

```

1 GAACCCAGTT GCTTCAGCGA GTCGAACTAC AGTTTTAACC TCATCAAATA
51 TGGCATCTCC CTTGCTTGCT GCAGCAGGGA TGGAGAAAT GTCACCTTCT
101 TTTTAAGCTA GCAAGCTTTT TCTTTTCTT TTTCTTCTT TATTTAAAAA
151 TTCTAATCAT GGATGCTTCT TCCGACCCTT ATTTGCCTTA TGACGGGGGA
201 GGAGACAATA TTCCCTGAG GGAATTACAT AAAAGAGGAA CTCATTATAC
251 AATGACAAAT GGAGGCAGCA TTAACAGTTC TACACATTTA CTGGATCTTT
301 TGGATGAACC AATTCCAGGT GTTGGTACAT ATGATGATTT CCATACTATT
351 GATTGGGTGC GAGAAAAATG TAAAGACAGA GAAAGGCATA GACGGATCAA
401 CAGCAAAAAG AAAGAATCAG CATGGGAAAT GACAAAAAGT TTGTATGATG
451 CGTGGTCAGG ATGGCTAGTA GTAACACTAA CAGGATTGGC ATCAGGGGCA
501 CTGGCCGAT TAATAGACAT TGCTGCCGAT TGGATGACTG ACCTAAAGGA
551 GGGCATTTGC CTTAGTGCCT TGTGGTACAA CCACGAACAG TGCTGTTGGG
601 GATCTAATGA AACAACTTT GAAGAGAGGG ATAAATGTCC ACAGTGGAAA
651 ACATGGGCAG AATTAATCAT AGGTCAAGCA GAGGGTCCTG GTTCTTATAT
701 CATGAACATC ATAATGTACA TCTTCTGGGC CTTGAGTTTT GCCTTTCTTG
751 CAGTTTCCCT GGTAAAGGTA TTTGCTCCAT ATGCCTGTGG CTCTGGAATT
801 CCAGAGATTA AAACATATTT AAGTGGATT C ATCATCAGAG GTTACTTGGG
851 AAAATGGACT TTAATGATTA AAACCATCAC ATTAGTCCTG GCTGTGGCAT
901 CAGGTTTGAG TTTAGGAAAA GAAGGTCCCC TGGTACATGT TGCCTGTTGC
951 TGCGGAAATA TCTTTTCCTA CCTCTTTCCA AAGTATAGCA CAAACGAAGC
1001 TAAAAAAGG GAGGTGCTAT CAGCTGCCTC AGCTGCAGGG GTTTCTGTAG
1051 CTTTGGGTGC ACCAATTGGA GGAGTTCTTT TTAGCCTGGA AGAGGTTAGC
1101 TATTATTTTC CTCTCAAAAC TTTATGGAGA TCATTTTTTG CTGCTTTAGT
1151 GGCTGCATTT GTTTTGAGGT CCATCAATCC ATTTGGTAAC AGCCGTCTGG
1201 TCCTTTTTTA TGTGGAGTAT CATAACCAT GGTACCTTTT TGAAGTGTG
1251 CCTTTTATTC TTCTAGGGGT ATTTGGAGGG CTTTGGGGAG CCTTTTTCAT
1301 TAGGGCAAAT ATTGCCTGGT GTCGTCGACG CAAGTCCACG AAATTTGGAA
1351 AGTATCCCGT TCTGGAAGTC ATTATGTTG CAGCCATTAC TGCTGTGATA
1401 GCCTTCCCTA ATCCATACAC TAGGCTAAAC ACCAGTGAAC TGATCAAAGA
1451 GCTTTTACAC GACTGTGGTC CCCTGGAATC CTCTTCTCTT TGTGACTACA
1501 GAAATGACAT GAATGCCAGT AAAATGTGCG ATGACATTCC TGATCGTCCA
1551 GCAGGCATTG GAGTATATTC AGCTATATGG CAGTTATGCC TGGCACTCAT
1601 ATTTAAATC ATAATGACAG TATTCACTTT TGGCATCAAG GTTCCATCAG
1651 GCTTGTTCAT CCCAGCATG GCCATTGGAG CGATCGCAGG AAGGATTGTG
1701 GGGATTGCGG TGGAGCAGCT TGCCTACTAT CACCACGACT GGTTTATCTT
1751 TAAGGAGTGG TGTGAGGTGG GGGCTGATTG CATTACACCT GGCCTTTATG
1801 CCATGGTTGG TGCTGCTGCA TGCTTAGGTG GTGTGACAAG AATGACTGTC
1851 TCCCTGGTGG TTATTGTTTT TGAGCTTACT GGAGGCTTGG AATATATGTG
1901 TCCCCTTATG GCTGCAGTCA TGACCAGTAA ATGGGTGGA GATGCCTTTG
1951 GCAGGGAAG CATTTATGAA GCACACATCC GATTAAATGG ATACCCTTTC
2001 TTGGATGCAA AGAAGAATT CACTCATACC ACCCTGGCTG CTGACGTTAT
2051 GAGACCTCGA AGGAATGATC CTCCCTTAGC TGTCTGACA CAGGACAATA
2101 TGACAGTGGA TGATATAGAA AACATGATTA ATGAAACCAG CTACAATGGA
2151 TTTCTGTGTA TAATGTCAAA AGAATCTCAG AGATTAGTGG GATTTGCCCT
2201 CAGAAGAGAC CTGACAATTG CAATAGAAAG TGCCAGGAAA AAACAAGAAG
2251 GTATCGTTGG CAGTTCTCGG GTGTGTTTTG CACAGCACAC CCCATCTCTT
2301 CCAGCAGAAA GTCCTCGGCC ATTGAAGCTT CGAAGCATTC TTGACATGAG
2351 CCCTTTTACA GTGACAGACC ACACCCCAAT GGAGATTGTG GTGGATATTT
2401 TCCGAAAGCT GGGACTGAGG CAGTGCCTTG TAACTCACAA TGGGCGCCTC
2451 CTTGGCATT TAACAAAAAA AGATATCCTC CGGCATATGG CCCAGACGGC
2501 AAACCAAGAC CCCGCTTCAA TAATGTTCAA CTGAATCTCA CAGATGAGGA
2551 GAGAGAAGAA ACGGAAGAGG AAGTTTATTT GTTGAATAGC ACAACTCTTT
2601 AACCTGAGGG AGTCATCTAC TTTTTTTTCC TCCTTTACAA AAAAAGAAAG
2651 GAAATATAAA AACCGGGTTT TTGCAACATG GTTTGCAAAAT AATGCTGGTG
2701 GAATGGAGGA GTTGTTTGGG GAGGGAAAGG AGAGAGAAGG AAAGGAGTGA
2751 GGTATTTCCC GTCTAACAGA AAGCAGCGTA TCAACTCCTA TTGTTCTGCA
2801 CTGGATGCAT TCAGCTGAGG ATGTGCCTGA TAGTGCAGGC TTGCGCCTCA
2851 ACAGAGATGA CAGCAGAGTC CTCGAGCACC TGGCCTGTTG CTCCAACATT
2901 GCAAAGACAC ATTATCAGTC CCTATTTCTA GAGGGATTAC TTTGAATTGA
2951 GCCATCTATA AAACGCAAG GTCTTGCCTT TTTTTTTAAT CAAAAGTGT
3001 CTGTTTAATT CATGAATTGT ATAGTTAAGC ATTACCTTTC TACATTCCAG
3051 AAGAGCCTTT ATTTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTACTG
3101 AGCTGTAACA AAGCCTCTTT AAATCGGTGT ATCCTTTTGA AGCAGTCTCT

```

FIGURE 1, page 1 of 3

```

3151 TCTCATATTG AGATGTACTG TGATTTTACT GAGGTTTCAT CACAAGAAGG
3201 GAGTGTTTCT TGTGCCATTA ACCATGTAGT TTGTACCATC ACTAAATGCT
3251 TGGAACAGTA CACATGCACC ACAACAAAGG CTCATCAAAC AGGTAAAGTC
3301 TCGAAGGAAG CGAGAACGAA ATCTCTCATT GTGTGCCGTG TGGCTCAAAA
3351 CCGAAAACAA TGAAGCTTGG TTTTAAAGGA TAAAGTTTTC TTTTGTGTTT
3401 TCCTCTCAGA CTTTATGGAT AATGTGACCG GGTCTTATGC AAATTTCTA
3451 TTTCTAAAAC TACTACTATG ATATACAAGT GCTGTTGAGC ATAATTAAAT
3501 AAAATGCTGC TGCTTGACA GTAAAGAGAA AAAAAAAAAA AAAAAAAAAA
3551 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3601 AAAAAAAAAA AAAAAAAAAA AAAAA (SEQ ID NO:1)

```

FEATURES:

```

5'UTR:      1-158
Start Codon: 159
Stop Codon:  2532
3'UTR:      2535

```

HOMOLOGOUS PROTEINS:**Top BLAST Hits:**

			Score	E
CRA 18000005109762	/altid=gi 2599548	/def=gb AAB95161.1 (AF029...	1575	0.0
CRA 18000005109763	/altid=gi 2599550	/def=gb AAB95162.1 (AF029...	1573	0.0
CRA 18000005227216	/altid=gi 4762023	/def=gb AAD29440.1 AF14277...	1572	0.0
CRA 18000004989660	/altid=gi 4502869	/def=ref NP_001820.1 chlo...	1570	0.0
CRA 18000005231972	/altid=gi 8134363	/def=sp Q9R279 CLC3_CAVPO ...	1561	0.0
CRA 18000004989700	/altid=gi 6680948	/def=ref NP_031737.1 chlo...	1560	0.0
CRA 18000004978791	/altid=gi 1705905	/def=sp P51792 CLC3_RAT CH...	1560	0.0
CRA 1000685681515	/altid=gi 6634696	/def=emb CAA71072.2 (Y0994...	1559	0.0
CRA 18000004989661	/altid=gi 1705903	/def=sp P51790 CLC3_HUMAN ...	1558	0.0
CRA 18000005226296	/altid=gi 4753144	/def=gb AAB88634.2 (U8346...	1556	0.0

EST:

		Score	E
gi 10993825	/dataset=dbest /taxon=96...	1562	0.0
gi 10934924	/dataset=dbest /taxon=96...	1336	0.0
gi 10952244	/dataset=dbest /taxon=96...	1251	0.0
gi 12383593	/dataset=dbest /taxon=96...	1205	0.0
gi 6591096	/dataset=dbest /taxon=9606 ...	1170	0.0
gi 10251711	/dataset=dbest /taxon=96...	1104	0.0
gi 2321385	/dataset=dbest /taxon=9606 ...	1045	0.0
gi 5594360	/dataset=dbest /taxon=9606 ...	975	0.0
gi 5422132	/dataset=dbest /taxon=9606 ...	965	0.0
gi 10327969	/dataset=dbest /taxon=96...	963	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:**library source:**

gi 10993825	Neuronal precursor cells-teratocarcinoma
gi 10934924	Whole embryo-mainly head
gi 10952244	Neuronal precursor cells-teratocarcinoma
gi 12383593	Small intestine-duodenal adenocarcinoma
gi 6591096	Lung-small cell carcinoma
gi 10251711	Breast-normal
gi 2321385	Schwannoma tumor
gi 5594360	Brain-tumor
gi 5422132	Testis
gi 10327969	Lung-large cell carcinoma

```

1 MDASSDPYLP YDGGGDNIP L RELHKG RTHY TMTNGGSINS STHLLDLLDE
51 PIPGVGT YDD FHTIDWVREK CKDRERH RRI NSKKKESAW E MTKSLYDAWS
101 GWLVVTLTGL ASGALAGLID IAADWMTDLK EGICLSALWY NHEQCCWGSN
151 ETTFEERDKC PQWKTWAELI IGQAEGPGSY IMNYIMYIFW ALSFAFLAVS
201 LVKVFAFYAC GSGIPEIKTI LSGFIIRGYL GKWTLMIKTI TLVLAVASGL
251 SLGKEGPLVH VACCCGNIFS YLFPKYSTNE AKKREVL SAA SAAGVSVAFG
301 APIGGVLFSL EEVSYYFPLK TLWRSFFAAL VAAFVLR SIN PFGNSRLVLF
351 YVEYHTPWYL FELFPFILLG VFGGLWGAF F IRANIAWCRR RKSTKFGKYP
401 VLEVIIVAAI TAVIAFPNPY TRLNTSELIK ELFTDCGPLE SSSLCDYRND
451 MNASKIVDDI PDRPAGIGVY SAIWQLCLAL IFKIIMTVFT FGIKVP SGLF
501 IPSMAIGAIA GRIVGIAVEQ LAYYHHDWFI FKEWCEVGAD CITPGLYAMV
551 GAAACLGGVT RMTVSLVVIV FELTGGLEYI VPLMAAVMTS KWVGDAFGRE
601 GIYEAHIRLN GYPFLDAKEE FTHTTLAADV MRPRRNDPPL AVLTQDNMTV
651 DDIENTMNET SYNGFPVIMS KESQRLVGFA LRRDLTIAIE SARKKQEGIV
701 GSSRVCFAQH TPSLPAESPR PLKLSILDM SPFTVTDHTP MEIVVDIFRK
751 LGLRQCLVTH NGRLLGIITK KDILRHMAQT ANQDPASIMF N (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 5

```

1      90-93 NETT
2     364-367 NTSE
3     392-395 NASK
4     587-590 NMTV
5     598-601 NETS

```

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

```

1      24-27 KKE S
2     330-333 RRKS
3     331-334 RKST

```

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 8

```

1      22-24 SKK
2     333-335 STK
3     529-531 TSK
4     613-615 SQR
5     631-633 SAR
6     642-644 SSR
7     658-660 SPR
8     709-711 TTK

```

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 13

1	27-30	SAWE
2	34-37	SLYD
3	92-95	TTFE
4	93-96	TFEE
5	105-108	TWAE
6	217-220	STNE
7	249-252	SLEE
8	383-386	SLCD
9	589-592	TVDD
10	666-669	SILD
11	674-677	TVTD
12	679-682	TPME
13	709-712	TKKD

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 18

1	49-54	GLASGA
2	53-58	GALAGL
3	72-77	GICLSA
4	88-93	GSNETT
5	189-194	GLSLGK
6	206-211	GNIFSY
7	234-239	GVSVAF
8	240-245	GAPIGG
9	245-250	GVLFSL
10	310-315	GVFGGL
11	313-318	GGLWGA
12	314-319	GLWGAF
13	408-413	GVYSAI
14	447-452	GAIAGR
15	491-496	GAAACL
16	541-546	GIYEAH
17	638-643	GIVGSS
18	692-697	GLRQCL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	99	119	1.810	Certain
2	182	202	2.131	Certain
3	233	253	1.398	Certain
4	256	276	1.019	Certain
5	290	310	1.770	Certain
6	321	341	0.797	Putative
7	361	381	2.093	Certain
8	400	420	1.539	Certain
9	473	493	1.739	Certain
10	496	516	1.218	Certain
11	540	560	1.568	Certain
12	570	590	0.975	Putative

BLAST Alignment to Top Hit:

```
>CRA|18000005109762 /altid=gi|2599548 /def=gb|AAB95161.1| (AF029346)
      chloride channel protein 3 [Homo sapiens] /org=Homo
      sapiens /taxon=9606 /dataset=nraa /length=818
      Length = 818
```

Score = 1572 bits (4026), Expect = 0.0
Identities = 764/765 (99%), Positives = 764/765 (99%)

```
Query: 27  GTHYMTNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 86
           GTHYMTNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE
Sbjct: 54  GTHYMTNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 113

Query: 87  SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 146
           SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC
Sbjct: 114 SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 173

Query: 147 WGSNETTFEERDKCPQWKTWAEELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 206
           WGSNETTFEERDKCPQWKTWAEELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA
Sbjct: 174 WGSNETTFEERDKCPQWKTWAEELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 233

Query: 207 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 266
           PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG
Sbjct: 234 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 293

Query: 267 NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLIEVSYYFPLKTLWRSF 326
           NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLIEVSYYFPLKTLWRSF
Sbjct: 294 NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLIEVSYYFPLKTLWRSF 353

Query: 327 FAALVAAFLVRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA 386
           FAALVAAFLVRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA
Sbjct: 354 FAALVAAFLVRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA 413

Query: 387 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 446
           WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD
Sbjct: 414 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 473

Query: 447 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMAI 506
           YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMAI
Sbjct: 474 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMAI 533

Query: 507 GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 566
           GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL
Sbjct: 534 GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 593

Query: 567 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 626
           VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL
Sbjct: 594 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 653

Query: 627 AADVMRPLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRDLT 686
           AADVMRPLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRDLT
Sbjct: 654 AADVMRPLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRDLT 713

Query: 687 IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDTHTPMEIVVD 746
           IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDTHTPMEIVVD
Sbjct: 714 IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDTHTPMEIVVD 773

Query: 747 IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 791
           IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN
Sbjct: 774 IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 818 (SEQ ID NO:4)
```

```

>CRA|18000004989660 /altid=gi|4502869 /def=ref|NP_001820.1| chloride
  channel 3; ClC-3 [Homo sapiens] /org=Homo sapiens
  /taxon=9606 /dataset=nraa /length=820
  Length = 820

Score = 1567 bits (4013), Expect = 0.0
Identities = 764/767 (99%), Positives = 764/767 (99%), Gaps = 2/767 (0%)

Query: 27  GTHYTMNNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 86
          GTHYTMNNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE
Sbjct: 54  GTHYTMNNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 113

Query: 87  SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 146
          SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC
Sbjct: 114 SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 173

Query: 147 WGSNETTFEERDKCPQWKTWAEIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 206
          WGSNETTFEERDKCPQWKTWAEIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA
Sbjct: 174 WGSNETTFEERDKCPQWKTWAEIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 233

Query: 207 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 266
          PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG
Sbjct: 234 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 293

Query: 267 NIFSYLEFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLLEEVSYFFPLKTLWRSF 326
          NIFSYLEFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLLEEVSYFFPLKTLWRSF
Sbjct: 294 NIFSYLEFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLLEEVSYFFPLKTLWRSF 353

Query: 327 FAALVAAFVLR SINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGFAFFIRANIA 386
          FAALVAAFVLR SINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGFAFFIRANIA
Sbjct: 354 FAALVAAFVLR SINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGFAFFIRANIA 413

Query: 387 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNNTSELIKELFTDCGPLESSSLCD 446
          WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNNTSELIKELFTDCGPLESSSLCD
Sbjct: 414 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNNTSELIKELFTDCGPLESSSLCD 473

Query: 447 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMAI 506
          YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMAI
Sbjct: 474 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMAI 533

Query: 507 GAIAGRIVGIAVEQLAYYHHDWFIKWECEVGADCITPGLYAMVGAAACLGGVTRMTVSL 566
          GAIAGRIVGIAVEQLAYYHHDWFIKWECEVGADCITPGLYAMVGAAACLGGVTRMTVSL
Sbjct: 534 GAIAGRIVGIAVEQLAYYHHDWFIKWECEVGADCITPGLYAMVGAAACLGGVTRMTVSL 593

Query: 567 VVIVFELTGGLLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE--EFTHT 624
          VVIVFELTGGLLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE EFTHT
Sbjct: 594 VVIVFELTGGLLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHT 653

Query: 625 TLAADV MRPLRNDPPLAVLTQDNMTVDDIENMINETS YNGFPVIMSKESQRLVGFALRRD 684
          TLAADV MRP RNDPPLAVLTQDNMTVDDIENMINETS YNGFPVIMSKESQRLVGFALRRD
Sbjct: 654 TLAADV MRPRNDPPLAVLTQDNMTVDDIENMINETS YNGFPVIMSKESQRLVGFALRRD 713

Query: 685 LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPFVTVDHTPMEIV 744
          LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPFVTVDHTPMEIV
Sbjct: 714 LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPFVTVDHTPMEIV 773

Query: 745 VDFIRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 791
          VDFIRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN
Sbjct: 774 VDFIRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 820 (SEQ ID NO:5)

```

FIGURE 2, page 4 of 5

```
>CRA|1000685681515 /altid=gi|6634696 /def=emb|CAA71072.2| (Y09941)
    putative chloride channel ClC-3 [Xenopus laevis]
    /org=Xenopus laevis /taxon=8355 /dataset=nraa
    /length=791
    Length = 791
```

```
Score = 1559 bits (3993), Expect = 0.0
Identities = 745/791 (94%), Positives = 771/791 (97%)
```

```
Query: 1 MDASSDPYLPYDGGGDNIPRLHKGTHYTMNGGSINSSTHLLDLLDEPIPGVGTYYDD 60
          MD SSDPYLPYDGGGDNIPRL+LHKGTHYT+TNGG+INS+THLLDLLDEPIPGVGTYYDD
Sbjct: 1 MDISSDPYLPYDGGGDNIPRLDLHKGTHYTVTNGGAINSTTHLLDLLDEPIPGVGTYYDD 60
(SEQ ID NO:6)
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
CE00039	CE00039 chloride_channel	1671.9	0	1
CE00420	E00420 CLC	1288.1	0	2
PF00654	Voltage gated chloride channels	1172.4	0	1
PF00571	CBS domain	78.1	7e-20	2

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00654	1/1	71	622 ..	1	621 []	1172.4	0
PF00571	1/2	645	690 ..	11	54 .]	31.4	5.8e-07
CE00420	1/2	32	697 ..	1	729 [.	1174.4	0
PF00571	2/2	726	778 ..	1	54 []	47.4	2.2e-11
CE00420	2/2	722	791 .]	867	942 ..	110.6	6.5e-32
CE00039	1/1	60	791 .]	1	804 []	1671.9	0

1 AATTCTATAC AAATATAATT ATATAGATAT ATATTACATA TACACACAAT
51 TGTTTTATCTT TAAAAATAAT TCAAATATGG CTACAAAAC TTTACAATAT
101 GAAGCATTGT CAGTATTTAT TTTACCGGGA GGATTTCCTT CATCAGTGAG
151 TGCTGACTGT CATTTTTCATT CTTTATGATC AAGTTGTAGA TCAGGAAAAA
201 CAAGTTAAGA GAGTGCCTAC AAATACCGGG AAAACTTGTG GATAGATTTT
251 CATTTTTTTAT GTAAAGACAT ATAAGAACAT GAATGGTATA AAAACAAAAT
301 CCTTTATAAA TGCCATACAA TTATATATTT AGAAAAATTA TATGGTGGTA
351 AAACATATAA AAGAACCACA CACTCCCAA TTTACATTGA GCTAATTTAG
401 TACAGTTAGC CTTTGTCAAA GCTTTCCTTG TTTAAAAAAA CTATTGGCTC
451 AGTGTGCAGG AAGGAGCATA GGAGAAAAA TTGCCAAGAA TATTTGAAAA
501 ATACAGAAAA TAAAGAAAAA AATCACCTAC TATCCTATCA AAAATTTTAA
551 TAGCTAGAAT CAGGATAAGA TAGAATATTC CTGTGGCAGT AATTCTAGTC
601 TATATTCTCT TCCTGGAACC CTGTCTCCA AATTCAGGT GAGATTTTAT
651 AAGAAGCTCT GTTTATCTGA GATTTAAAT ATAAAAACTT GATTTAACCT
701 ATACAGTTTT TTA AAAAGAC CCTAAATAAG TAAAAATTAG TACTCCACAA
751 ATTGAAGAGA ATTTCTCTCT TCTCTTTACT GCCCTCTGAG TTTTCTCTTT
801 CCTTCTCTCA CCTCAATTT TCATGTAAAC ACTTTCAGTT CGAGTGGACC
851 TTAGAGATTG TCTCATTCAA TACTTTAGGA AAACAAATTT TATAGAACCC
901 TTGAGTTCTG TGGAAATGCT TCTAATGAAC AACACCTTTT GTTGTGTGTG
951 TTGTTTAGTG AACTGTGTGA ACAGGCATTT CAGGAGGAGA ATCTCCCAGT
1001 CTAGAGGAAT CCTCTCAGAG GTAGCTATAA AATATTGAAC TCTGATCTTC
1051 AATAAGCATT GTGCGGTTTT TGTTTTTGTT TTTAATGACA GTTTTAAACA
1101 AGAAAGTTCG TTTATTTCTG AACTTCATAA AAATTTCTAT TAAAGAGACA
1151 ATTTCTGAAT TTTATAACAA TTTCTAGAAC AGTTGAGTAC CTCACTTTGA
1201 GACACATTTT TGCTAAAAGT TAAAAACACA AAACCCTTAT GAGATAAAAT
1251 AGGAAGCTAG TAGAGATAGG AAAGTCTCT GCTTAGTAAA CCTCTTTTTT
1301 GCGTAGTTTA GACACATACA ATAGTAAAGT TACTTAGTAC GTTGATAGTT
1351 TTCTTTCTCC TCAAAAGCTA CAATGCTCTA CTAGCTAGTT CCTTCAAGAA
1401 AGGAAACAGG AAGCCGCTGG AGGAGATTGG TGAGTGGGAT AAAACACTAT
1451 TCAACTCTTC AGTTATTTCGG TTTTAAATC CTCAATGAAA GGCTGCTGTA
1501 TTATAGAGTA TTTTTTTTTT ATTTTAAATA GACTTAGAAC CAAGTTTCTT
1551 GAGAAACCTT TGGCATATTG TAGTTTTTTT ATGGCTATGA CTCACATGAC
1601 ATTACTGTAT AAAACTAGTA CATTCTCTCG TAAACCACA CAAACTTACT
1651 AGAGTGCTGC TCTCATTTTT CTACATTAGA AATGAAAAAG GGCATTGTCT
1701 GCATCAAAA TTTCTTTTTT ACATCTCTGT ATTACTTTTT CCCCTTTATA
1751 TTTATCTTAA AACCAAAAGA AATAATGTTT CTATGTGTTT ACTGTAGTTA
1801 CCACTGATGC TACCGAAGCT GTATTGTGAG TGTTCAAAA TTCTCAAACC
1851 AGTTTTGTGT GTTGTACTTG GAGCTTAGTC ATTGTCATAC GTAGCAGGAC
1901 CTGATTAGA AGGCTGTGCC GCCTCTAAGC CTGCTAGAT GTAGCCACT
1951 AGCAACCAGG CTGCAATAAT TTCCCTTTGA TGACATCATC CACTGTGGAA
2001 GAACCCAGTT GCTTCAGCGA GTCGAACTAC AGTTTAAACC TCATCAAATA
2051 TGGCATCTCC TTTGCTTGCT GCAGCAGGGA TGGAAAGAAAT GTCACTTTCT
2101 TTTAAGCTA GCAAGCTTTT TCTTTTTCTT TTTCTTCTTC TATTTAAAAA
2151 TTCTAATCAT GGATGCTTCT TCCGACCCTT ATTTGCCTTA TGACGGGGGA
2201 GGAGACAATA TTCCCTGAG GGAATTACAT AAAAGAGGTA ATACTATCCC
2251 CTTGCTGTGA ATTCTCTGTT GGTATGTTT GCATGCGGCT GGGCGGTCTT
2301 CTAGCTTAAA CTGGTCTCTG TTTGTCTTT AAATACTGCA GTACGTTGTT
2351 TAGTTGCCCT GGGTTGTTAG TAAGGGGAAA ATGCAACCTT CTGAATGGTT
2401 GTGTAGCCAT CCCTGATTGT TTTCTCTGTG CAGATTAGTA CTGCTTCAGA
2451 TCACGTCGGG CTCCGACTCC ATCTTCTGCA TGAAAATCTT CTTTCTAACT
2501 CTGAAAATGA ATTAATCTGC TTTTACAGCC AACTAAAGTC GTGTTGGTTG
2551 GCATCTAAAA AGTAATGTTT TTCTTCCTTC AGAAAACTTA CATTTCCTTT
2601 AATTTACACA GAGAAATCAG GTGCCTATGT ACCATTATAT TTAGCTGCT
2651 GCCAATTACC ATGTAGATTT TACACCACAA AGTAAATTTA TAGCAAAAGC
2701 TTTACCTACA TTTTAGAACA TTTTAAATG ATAGTAAAGA TGAATAATTT
2751 CTATATTAAT ACTTTTTATT TAATATGTAT TTCGGCTGAG TAACATACTA
2801 CATTGTCTTC CACAGGTATC TTGTGAAAT TGATATGATA AAACACATTT
2851 GACTAAATGT CAGAAAAAAT AATATTGGTT TGTGAAAAGC AGAAGAGCAC
2901 CCAGCATGCC TGTAAATCTT TTGGCAGGCA CTTCTCAGT CTCCTTAAAA
2951 TTAATTGCAT GTTAATTACT ACCCTTTTTT TCATTTTTGT TTAATTGCTT
3001 ATTCGAAAAA CAGACTGGTC GACATTTGTT GTCCTAGAAA AAAATTGAAC
3051 TTCAAGAAAA ATCTCTTAGC TTATGTGACT TCATTTTGA GCCACATTAG
3101 TTTGAATTAC TGCAATGATAT TATAAACTCA CCTTATGATT TAACCCAAAC

FIGURE 3, page 1 of 27

3151 TTTTATTTGT AAGTATATAA GGAAGTAATA ATGTTTTTCT AATATAATTA
3201 GCCTGCTTTA TTTAAAATAT ACTTTGTGTT CTGATAACAC TTTTTTTTAA
3251 GTATTAGGTT CCACATAAAT TTAAACATTA TAATGTATTC AACAAATGTC
3301 TGTGGGTTGC ATTTGTGTCTG CTACACACTA TTTTAGGGTC TGAACAGTTG
3351 TAGCATTATT TATCTTGCAG TATTCTGTAG TTAGTAAAAA CTTGCTTTTT
3401 ACATTTTGAG AAAAGCTGTG TAAGGATCAT GTTACATACA TTGTGCTTTC
3451 TCTTACAGAG TTACCTTCTT AATAAAATTT TGATATATGT GTATATGTAT
3501 ATGTTAGAAC ATTTGGAAGA AATATCTAAA AGCATAAAGA AGAAAAATAT
3551 TTCTTGTAAT CACACCACCC AGAGCTTTTT AAATTTTTTT TCTTAATGTT
3601 ACGATCATAA ATTCTTCTAT TTCCTATGTT CTGATTATCA GTTTTCTGGT
3651 AAGGAGTTCT TTAAACAGGA AGCAAGGTGA ATGAATAGTG ACTGTTCAAA
3701 TGTCACATTA TTTGCTAATC AGTAATTAAA CTGTAAAACA AGACAGACTG
3751 TATTTTCCTC ATGCTATTAC AACATTTGGT TGTAAATGAT GATAGATCAG
3801 AATACCTGGG CTTCAGAAAT TTAAATTCCT TTTGTGAAGC TTAACAGTCT
3851 TTGACAGAAC TTACTTATGG ACTGTCTTAG TGTAATAATAT GCAAATAATA
3901 AGAAATAAGT CAAAACCTAT GTGAGAGTAG GCATGGTTAC TGATATTACC
3951 TAAACGTAAG CTTTTTATTT CTATTATACT TTCATAAATA ATCCTTTAAG
4001 AATCTTGCTT AGGATCTAAA TCAGTCCCAC TCTTGGCAGC TCAAATAGGT
4051 TCTTTATCCC TTGATGAGAC TTATTCTATT AATATAAGTC ATTGTTATTT
4101 GAAAGTAACA TTGTGTATGT GTAGTAGAGA TAAGTCAGTT ATTAGGCTTT
4151 CGTGACTGTA CTGTATTACC TCAAACATAC TGATGTATCC TAGTGTCTAT
4201 GCGTAAGATG TTATTTTTTG TCCATAATTT ATGACCTGTT GTAGCCATGG
4251 GTCAACACAA TGGAATTGAT GGAGACAGGC AGCTAACAAA TCGAAAAAAC
4301 TGAATCAGCT TCCCTGTGAG GAAGAACAAA ACTATAATGA TTAATAATGA
4351 TCTTCAGCCT GATAGTGAAG AGGCAGATAA AGTATAAAAT TGTGAAGGAT
4401 ATCAATAAAG TAAACATGGA TCTGTTTAGT AAATCCCTGA GTGCTATAGC
4451 CAAGGATTAC CTTTGTGAG TAAATGAAT TTAATACTAC TTTTCAAGGC
4501 GAGATGGTAA ATGGTGAAGC TTCCTATTTA AGTAAATAAT GTCAAGTCTG
4551 GAAGTATAAG TAGATTCAAA TTAGAATTAG TTTGATATAC TATTGATAGA
4601 TTAGAAATTA AGATGACATT TCAGAAATAG CCATCTTTAG GGGTAGATTT
4651 CCTATATAGA AACAATCAAG CTCTCTCAAA ATGTCTCTTC CTTTTTTATC
4701 AGGAAAAAAG ACTTGGCTTA TCTGGACTGT TAGTTTACA CTTTTTCTTC
4751 TTAATTTGTT CAAGATGTTT AAGTAGTTT AGAGGTCAAA TTTCTTTCTT
4801 CTACCAACCC TTTATAATGG ATTTGATTCT TTTGGGCCTG AGCCTCCATT
4851 TACTCCATGA GGGGCCTTTA ACAATTATTT AAATNNNNNN NNNNNNNNNN
4901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5101 NNNNNNNNNN NNNNNAAAAT AGTAATATTA ATAATAGTTA ATATTTATTA
5151 GAATTCCTG TTAGCTGGAT ACTGTCCCTA AGTGGGTTTT TTTGTTGTTG
5201 TTGTTGTTGT TGTGTTTTTC TTAAGAGAGA GGTATCACTT TTTACCCAG
5251 GCTGGAGTGC AGTGGAGTGA TTATAGCAAA TGCAGCCTTG AACTACTGGG
5301 CTTAGATCCT CCGTCTCACC CTCCTTGGTA CCTGGGACTG CAGGCTTGCA
5351 ACACCTTGCC TGGCTAATTT AAAAAACAAA ATTTTTTTTT TTTTAGGGA
5401 GAGTCTCACT ATGTTGTCCA GGCTGGTCTC CAACTCCTGG GCTCAAGCAA
5451 CCCTCCTGCC TTGGCCTCCC AAGTAGCTGA GATTACAGGT GCGAGCCACT
5501 GTGCCTGGCT TGTCTAAGT GCTTTATGTG TATGAAATTA TTTAAATCCT
5551 CATCACAAGT TTATGAAGTA GGTACTGTTA TAATCCCAT TTTCTAGTTG
5601 ACAAGACTGA GGTAAAGGAAT TGTTAAGGAA AAGTCAGAAT TCCATCCAGA
5651 TATTTGGCTC ATACTTTAAT CATGAGGCTA AACTGCTTCT CTCTACACGT
5701 ATCTTCATAG TAACTTGTGT TTTAAGTCTG GTAGAAGCAT AAGAAGTTTA
5751 AACACAGACA GAATCCTGTG GAAGTTAGTA AATTTCTAGT GAACGATAGA
5801 AATGATAGAA ATCTCTCTT CCCCCAAAGT CCAAGAACA GATTAGTCTG
5851 CTTTTGACAA GTGTTATCAA AGTAGACTGT TCTCACATAC ACGGGGGACT
5901 CAATAGGGCA TTCCTGGTGG ATATAATAAA ATGAGTAAAT GCGATAACAG
5951 GAGGAAATGC CTAGTGTGTT GCTCTTGGAT TAGTTTTGAT ACAACAAAGG
6001 CAGCTTTGTT GTGAGTCAGT AGAGAGGGTA GTGTAGAAAG GTGGAAGTTG
6051 GAAGAGTGGC AGATCCTAGA GGAATAATGA TGGGCTTAAA CCACAAAAG
6101 TGTCGCTTTG CCATTGAAAT AAAAGTTTGG GGTCTTATTT TTTCAATTTT
6151 CTCCCTGAAA TTATTTCTTG ACATTCATTA GCTCAGCAGT GTATCTAAAT
6201 AAAGCTTTTT TGGGTTTCTA TTATAATAGA GGTTGTTTCC TTTTCTTCC
6251 CTTTGAAGAG TATCATTTTT TGACATTAT TTGAAAATCC AGGTGTTATA

FIGURE 3, page 2 of 27

6301 TGATATTCTT ATTGCCAGAG GGACATTCTG CAGGCTCTTT GTAAAAATGAT
6351 TTTAGGATTC AGATACTTAT TATATTTTAA TTGGCCCTAA TATTTTATCC
6401 AACTAGAAAA TTAACCTCT TCTTAAAAAT TAATCCATCT AAGTGTCTGT
6451 AAATTAAAGG AACCACTAAA GATTCTTTAT TTGGTGTGAG AAACCTCTTG
6501 TTTCTACAAC AGTAGTATAA AACAAAGCCT GTTTTTAAAT GTACTTTTCC
6551 CACAGTATCT GAATTTCAAA TCTTCAATAA AATCTGGTTC ATATTACTAC
6601 CTCTAGCTTG ATTTTCTAAA AATAGCTGAC ACTTTTAGTAT GGTAAATTTT
6651 ATGCCATCTC ATGGCTTGTC AGAAATGCTT TGTATCAAGA TTCCGAGTGT
6701 TGAACAGATT TCCTGCCGCA TTGATTAAGT TTGTAATTTT GGCTATTTTC
6751 CCAGCATCGA GGTTCCTGCT TTGCGTTTAT GCAGGAGACT GGTAGTTTAA
6801 ATTGAACCTT AAGGTTTTGT TTCTTGTTTT TAAGTTAACA TATGTTTAAAT
6851 TTCTAGTTTC TTTGTAGCCC TTTGCAACTT TAATTAGGTC ATAAAAATGGA
6901 TTTACTCTAG TTTCTCTAAC AAATTTTATA AATTTATGAA ATATGAAATT
6951 TAGCAAATTT TATAAACCTT TTTATTTCATG TATTGTACAG CTCATCATAT
7001 TTGCAGACAT AATAATTGAA TGTGGAACCT GTTCCCAATT ACACAGATGT
7051 CTTAATATCC ACCTTATCAT CTCTAACTAA AGGATGTGGC TTTTATTTT
7101 TGAGGTGGCA ACAGAACAGA AAAGAAAACA GTGAATTGAG TAATGGGCTT
7151 AGTATTGCTG CTGCCTGGTT GTGTATCTTT GGTAAACTTC TTTGAGATTT
7201 GGCATTAACT TGCAAGTCTT TGCAGTTTAG ACAGTTAAAT ATGACTGAAT
7251 GGCTGAACAA ATTTTAATAG CGTATGCTTC TTTTTTGCTA TTTATTTACC
7301 CAGTAGACAT TTAATTGACC ACCTGCTAAA TGTGAGGCAC TATTCTTGCC
7351 ATTACCTTTT TAATCTTTGA TTTGGAGTCT GCTAACATTC TGGAACCTCC
7401 ACTATCAACT TAGAACGTTT ACTTTCCCAT CCCTTACCAG GATGGCCATT
7451 TCTTATCAGT AGGGTCACAG AGAGAGAAAA AAAAAACCAT CTGGGGCTAG
7501 ACTTCTGCT CTTAACATAC AGAAGCAAAAT AGGTTGTGAA GGAATACATA
7551 GTATTTTGGG TTTCTGCCTC TTCCTTCCAT AATTTTTTTA AAAAGGTTCA
7601 TATGTTTTAT GTGTGTCTTA TGTAAACAGTA ATCTGCATTA TGAACCTAAA
7651 TGACGAGGAT CACCATTTCA CATCTTTGGA GATTGATCAC AGAGGTAATA
7701 AGTAACTCTT TTTAAATAAC TATATGCATC ATTTTTCATG TAAACTATAT
7751 ATTTGGATAA ACCCCTTTGA GAAAAGGCTT AGGCTCCTGC CAGTGTCACT
7801 GTGATATTTA CTAATAAGCT CAGTTTAAGG CGCAGCAATT AAGGTTGTGT
7851 TGTTTTTTTT TTTTAAAGTT CAGTTCAGCA AATATATGTG GAAAGCTTGT
7901 GGGTAAATTT ATATTTGTAT TTTTGGGAAA GCAGACAATT TTATTAATGC
7951 CTATATTTTT CTAGTTCAGT GTTTGTCAAA CTCAAGTTT TAACATGTTG
8001 ATCATGAAAC CAGTTGACTT GTGACCAGTA TTTTAAAAGG AAAGATTAAA
8051 AAAACAAAAT AAAATATCAG TATATACCAA GTAGTAAGAG TAAGCATTGT
8101 TTTACTAACT TTGGTTTTAT TTAAGTACAT ATCTATATAC TATGTCACTG
8151 AGAAACATTT TTCCACTTCA TGTTTGAAAA ACATTTCAAA AGCTAAGAAA
8201 AAGTTTGAAA ACCTGTTTGT AAGTACACCT GGGGTAAAGG TACACCCTGT
8251 GGCATAAGAT GTCGGAACA ACTGAGGGTA AGAATGGGGA TGCATTACTA
8301 TCGTAAACTT CTGCTAAAGC ATAAGGATGT GAGTGTGGG AGCAAAGCAG
8351 TGCTCACCAC TTCTGCAATT TTCTATTGCA GCATTTTAAA TAATATGGGA
8401 AAAAGTGGAC TGCAACCAAA GGCAAGAGG GATGGTGATG GTGAAGGGTA
8451 AGATTGTATT TATTGTCCAA AGGCTAAGTG CATATACATA TGTGTTGGG
8501 AGAAGGCATC ACGTAATAGT TCTTAACCTA CTCTGAGAGA AGGTTGTCCA
8551 CATTTCTTAA AGTATACATG TAAACCAACA ATGAAATTAT TTTAGTGACT
8601 TGAGAAATCAA AGTGCTAGAG TTTGAATCCC TGTTCTACTA CTTGCTAGCG
8651 GTGTGACCTT GGGCCTGTTT AACTCTTGAC ACCTTGTTTT CCAAATTTAT
8701 AAAGTGGAGA TAATAATATC TGTCACATTG TGTGTTGTG AGGATTATAT
8751 GAACTAATAT ATGTAATGTC CTGAGAACAA TGTCTGGTAC ACATTAAGTT
8801 AATTAAAAAT AGCTGTTCTT ACTGTTATTA TTAGACATGA GCTAGATAAC
8851 AGTGGCCTCT ACATGTGAAA GATTATTTTA ATTCTGATGT AGTTCAGTTT
8901 ATCTATTTTT TTTATTTTTG TCCCTTTTGC ATTGATGTCA TATCTAAAAA
8951 ACCTGCCTAA CTGAGGATCA CAAAAATTTA CTCCTGTATT TTATAATTTT
9001 AGCTCTTTAG ATCTAGGATC CATTTTTCAT TAATTTTAT ATATGGTGTG
9051 AGGTAGGGGT ACGGTTTCAT TCTTTTGCAC GTGAATAGCC AGTTGTCCCA
9101 GCATCATTTA TTCAAAAGAC TATTCTTTCC TCACTAGAAA AAATATTTCT
9151 TTAAAGAATA ATGAATCCTT TTTTTTTTCT TTTTAAACGC TGTACTCAG
9201 TTGGAAAAAG AATAATGAAT AATTTTAAGT AATTTTCCTA CAGGTAAATT
9251 TAAGTCTTTA TGTTTAGATT ACACATATTA GGAAATAATG GATTTGTATT
9301 CCATAGGTAT GCTTGATCTT TATAAAGTTC CCTGTCTCTG GAAAACTAA
9351 AATAAGGCAA AACAATCTTC TTAGTAGAGT TATTTTACA AGAAAGTTGC
9401 AAGCCAGTTT TAGTTCATCG ATTGGATAAT TTTTCTGCT TGCTGGAGGT

FIGURE 3, page 3 of 27

9451 ATTCAGTAT TGGTAATACC TGAACATGA GGATGCATGA ATGATGCATT
9501 TTAGGAATTT GTTCTGTGT CCATACCAGG CATAATGAAT TAAGTTATCT
9551 GTTAAAAATA CAGGATTTTT GCTCAATATA CAGTTGTAGA AGAACTCATT
9601 GTCCAAATTT TTAAGACTTT TTTTCTTTT TTTTGTGAG ATGGATCTCG
9651 CTCTGTCGCC CAGGTTGGAG TGCAGTGGCA CAACCTCCAC TCACTGCAAC
9701 CTCCACCTCC AGGGTTCAAG TGATTCTGCT GCCTCAGCTT CCCGAGTAGC
9751 TGGGGACTAC AGGCATATGC CACTATGCCC GCCTGATTTT TTTTAGTAGA
9801 GATGGGGTTT CACCATATTG GCCAGGCTGC TCTTGAACCTC CTGACCTCGT
9851 GATCCACCCG CCTCAGCCTC CCAAAGTTCT GAGATTACAG GTGTGAGCCA
9901 CCGCGCCCGG CCAGACATTT TTTTTTTTTT TTTTTTTTTT GCTGTCTTTG
9951 TCATATTGTT AGTCTTTTGG TTAAGCGATA TTATAACTTA GTCATATGAG
10001 TAATATAATG CAACATGCTG AATTGTGTGT GTGAGAGGGG GTTGTTTTTT
10051 GTTTGTTATT TGTTTTTTAA ATAGAGATGA GATCTCACTG TGTTTCCCAG
10101 GCTCCCTTGA ACTCTGGGC TCAGATGATA TAGCCTCCTG CCACAGCGTC
10151 CTGATTAGCT GGGACTACAG GTGTGCACCA CTACACGTGG CTTTCCTGAT
10201 GAAATTTTAA ATACCCAAAT ATTTGAGCAG AAATAATAGC TTGTGTTTAT
10251 TGTTTTTCTA CTATCTGTCA AGTATAGTAT TAAATGTTTT ACATAATTTG
10301 TCTCCAGTCC ACATACAATA CTCTAGTAGA AGTGGGTAAC AAAACCAAGG
10351 TACTCAAAGA GGTAAATAAG TAACTTGCGC TGGATCACAG AACTAACGGG
10401 AGGCAGGGCT GGAATTTGAC TCTAGGTCTT TCTGACCTCA AAGTGCAGTA
10451 AAGTCATGGA ATTTCTCTAC TAGGCCACCT GGAAGAAAAG TGATCTTTTT
10501 TCCAGTCTTT TTTGTTACTG TTTTTCAGCC AGGAGATAGT AGAGTTAGGT
10551 AGTAGAATAG TAGTCACTGG CATCCGGTAG TCAGCCCTCC AAAAAAGTTT
10601 TTAGTTTTTT TTTTTTTTTT TGTCTTAAAC TTGGAAGCTA CTAACTTTCA
10651 GGTCATACTT TCTTATCATC CAAGAGCTGG ATATTTAGGT AGCAGAAACT
10701 ATGGAATTAT CCTAAGTCCT CTTGAAGCTT CAGCTGTTAA AATTAATTGG
10751 TTCTGATTAA CACTGTGCTC AAGATTTACA TTTCTAGGAG CCACAGTTTG
10801 ATTGGTCTAA CTTGGATCTA TGTGTTTTCT TTAGCTGGGG AGGAGAAGGT
10851 ATCTTGATTG ATACCTTCAC CAGGACTGCA TGCAGTGAGG GACAGAAGTT
10901 TCCTTAAAAA AATTGGGTTT TGTATAGGA AGAAGGGGAA GGAGATACCA
10951 AGTGGGCAAA ACAATCAGGT TCTATTACAT AAATAATAAA CCTAATGTGA
11001 CGATAATAAA TGGATAATAT GATTATTTTA AGTTTGAAA TATACCTGGT
11051 TATTAGTATT GGATATCTGG TAGTGGGGTT GGAGAAAAAG TCGAGAATAA
11101 GAAAAGACTT AAAATCGTAA AAATTAAGT GAAAAGAGGA TGGCTGAGCA
11151 GATACATATA TGTTAGATAA TGTTTATAAT GGCAAACCAA CCTGAAGATT
11201 TGTTTAAATT GTAGTATGTA GCCAGGTGTG GTGGTGCTTG CCTGCAGTCC
11251 CAACTACTTG GGAGGCTGAG GCAGGATGAT TGCTTGAGCC TAGGTTTGAG
11301 GCTACAGTGA GCTATGTTTC CACCACTGCT TTCCAGCCTA GGTGGCAGAG
11351 CAAGACCCCA TCTCTAAAAA AATAAAGTAA AATGAATAAA TTATAATATG
11401 TTATGACAAA TATAGTTATC TGAAGTCACA GAAATGTGC ATGTGCATTT
11451 AATGATGTGA AATAATTTTT AGGAAGTATG AATAAAAAAA TCAACTTTTA
11501 AGTGTGGCTA GTATGATCTT ACCTGTATCT CACTTATAGA AAATATAAAA
11551 GGCTGAAGCC AGTCACCAGT TTAATAGTTC TAACCTCTTG TTTACTTGAT
11601 TCCCTTTTTT CTCCTCCCCA GCAATCCTCA TATAGTTAGG TAAAGTTGGT
11651 TCTTCATCAG GCTTGTGCA GAAACCCCTA AGCCTTTTTA CTTAAAGCTT
11701 TTTGAAACCC AGAAACCCAT CTTTTGAATT CAAAAGTTTT GACTGTTATT
11751 AGTCTTTTTG TATGTTTGTT GGCCGCATAA ATGTCTCCTT TTTATGAACA
11801 GAGAAGTGTC TGTTAATATA CTTTGCCAC TTTTGTATGG GGTGTTTGT
11851 TTTTCTCTTG TACATTTGTT TAAGTTCCTT GTAGATTCTG GATATTAGAC
11901 CTATGTCAGA TGGATAGATT GCAAAAGTTT TCTCCCATTC TGTAGGTTGC
11951 TTGTTTCATC TGATGATAGT TTCTTTTACT GTGCAGAAGC TCTTTAGTTT
12001 AATTAGATCC TATTGTCTG TTTTGGCTTT TGTGCGCATT GCTTTTGGTG
12051 TTTCACTCAT GAAGTCTTTG CCAGTGCCTA TGTCTGAAT GGTATTGCCT
12101 AGGTTTTTCAT GGTTTTGGGT TTTACATTTA AGCCTCAAAT CGATCTTGAG
12151 TTAATTTTTG TATAAGGTGT AAGGAAGGGG TCCAGTTCCA GTTTTCTGCA
12201 TATGGATAGC CAGTTTTCCT AGCACCATT ATTAAATATTA AATAGGGAAT
12251 CCTTTCCCCA TTACTTGTTT TTGTCAAGTT TGCTGAAGAT CAGATGATTG
12301 TAGATGTGTG GTGTTATTTT TGAGGTCTTT GTTCTGTTCC GTTGGTCTGT
12351 ATATGTGTTT TGGTACCAGT ACTATGCTGT TTTGGTTACT GAGCCTTGTA
12401 GTATAGTTTG AAGTCAGGTA GTATGATGCC TCCAGCTTTG TTATTTTTCG
12451 TTAGGATTGT CTTGGCCATA CGGGCTCTTT TTTGGTTCCA TATGAAATTT
12501 AAAGTAGGTT TTTCTAATTT TGTGAGGAAA GTCAATGGTA GCTTGATGGG
12551 AATAGCGTTG AATCTATAAA TTAAGTCGGG CAGTATGGCC ATTTTCATGA

FIGURE 3, page 4 of 27

12601	TATTGATTCT	TCCTATCCAT	GAGCATGGAA	TGTTTTTCCA	TTTGTGTTG
12651	TCGTTTCTTA	TTTCCTTGGG	CAGTGGTTTG	TAGTTCTCCT	TGAACAGGTC
12701	CTTCACGTCT	CTTTTAAGTT	GTACTCATCA	TCACTGATCA	TTAGAGAAAT
12751	GAAAATCAAA	ACCACAATGA	GATGTCATCT	CATGCCAGTC	AAATGGTGAT
12801	TATTATAAAA	AGTCAAAAAA	GAATAGATGT	GGGTAAGGCT	GTGGAGAAAT
12851	AGGAATGCTT	TTACTACTGT	GGTGGGAGTG	TAAATTAGTT	CAACCATTGT
12901	GGAAGACAGT	ATGGCGATTG	CTCAAGGATC	TAGAACCAGA	AATACCATTT
12951	GACCCAGCAG	TCCCATTACT	GGGTGTATAC	CCAAAGGATT	ATAAATCATT
13001	CTGCTATAAA	GACACATGCA	CACGTATGTT	TATTATAGCA	CTATTTACAA
13051	TAGCAAAGAC	TTGAAACCAA	CCCAAAAAGC	CATCAATGGT	AGACTGGATA
13101	AAGAAAATGT	GGCACATATA	TACCATGGAA	TACTATNNNN	NNNNNNNNNN
13151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
13951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
14501	NNNNNTAAAAG	ATACATCCTT	TATTCATGCG	TAAGATGAAA	TCGAGAGGTG
14551	AAATTGGATA	TACTGTTGCT	TTTAAAAAAT	TTTAACATAT	ATGTAATTTT
14601	TTGTACTTAT	CTCATTTTAG	CCTATATAAG	TTATATATAT	TTTGTGTTGT
14651	TGTTTGTGTTG	TTTTGTTTGA	GATGGAGTCT	TGCTCTGTCA	CCCAGGCTAG
14701	AGTGCAGTGG	TGCAATCTCG	GCTCACTGCA	ACCTTCGCCT	CCTGCATTCA
14751	AGCGATTCTC	CTGCCTCAGC	CTCCTGAATA	GCTGGGATTA	CAGGCACCTG
14801	CCACCGCGCC	CAGCTAATTT	TTTTTATTTT	TAGTAGAGAC	AGGGTTTCAC
14851	CATCTTGGCC	AGGCTGGTCT	TGAACTCCTG	ACCTTGTGAT	CCATGTGCCT
14901	TAGCCTCCCA	AAGTGCTGGG	ATTACAAGCG	GGAGCCACCG	CGCCCCGGCTG
14951	TAAGTTATAT	CTTACACAAA	TCTAGGTTTC	ATTACAGAGAA	TTATATGCAA
15001	AGAAACAGTG	CAATAGGATT	ATTTTAAAGC	TATTGTTATT	GTTAGAAAAC
15051	ATAATACCTT	TAAAATTCCT	TTTCACATTA	GAAATATAGT	GGCTTCTCCC
15101	CAGTTTAGGA	TAGAAATTTT	CCTTTTCTTC	TCCTTCTTTA	TACTATTCAG
15151	ATTTGCATGT	TTGACAGAAC	AAATTATAAG	AGAAAATATT	TGAAATGTCA
15201	CATACTAAAG	TAAATGTTTG	AATGTTTGAA	AATTTTCTGG	TTTTTCAGAGA
15251	TTTTGAATTG	CTGAATCGTT	GTGTAAATTA	AGATGTTGAG	TAGTTTCCAC
15301	AGAGTAATTA	TTTGAAAGTC	ACTGAAAGCA	AGACACATGC	CTAATGTAAA
15351	TGTTTATTGC	ACTACTGTAC	CTTTTCTAC	CTCATAAAAA	TGAGAATAGC
15401	AGTCTGTAAT	TTTCCACTTC	GTCATTGTA	AGTCTTTGCA	GAAATTCATA
15451	TTTTGTTTGC	TTATTATCTT	CACGCTGTAA	ATAGCTTGAA	AATTCCTTAA
15501	GTGGGGCTAG	CGATGTATTA	TGGATACATG	TTAAGTGGTA	TAGAAATTC
15551	ACTTTTTTTT	TTTTGCATAA	AGAGTAACAA	GACCAAGTAGT	CCATATTTCT
15601	TCAGCTCTAC	CCAGAGAAGG	GCAATGTAGG	AGGGAAAATG	AAGTTTGCAA
15651	AATATTTTCAT	AGTAGGCTTT	TTCTTAAAGT	AACTTCAGAC	TTACAGAAGT
15701	TTAAAAATAG	TACAAAGAAT	CCCCATATAC	CTGTCACCCC	AATTCCTGAA

FIGURE 3, page 5 of 27

15751 ATATTAATAT TTTACCACAT TTGTTCATTA TGTCTGTATT CTCCAAGTAC
15801 GATATATGCC ATTATATGTA ATATGTAGCA TTTTATATAG ACATAGGGCA
15851 TGTATGCACCT ATATATTTTT TTCTGAGCCA CATAAAGAGT AAAACGCAGA
15901 CATGACGTGC TTTTACTCCT AAATACTTCA GTGTGTGTAT TCCCTCAAGA
15951 AAGGGCATT TCTTCTGTAT AGCTACCGTA CACTTCTACA CTTTTCAAAA
16001 TCAGAACATT TACATTGATA CCATACTATG ACATGATCTG CAGACCATT
16051 TCCAATATGC CAGTTGTCCC ACTGTGTCC TTAGTACAAA AGAAAAAAGT
16101 TTTTTTTCCT GGTCTAGGAG CTAATCCTGG AGCACATGTT ACATCCTGTT
16151 GTTTTAATCT AGAACCGTTC CTCAGTTCCT TATCTTTCAT AACCTTGACA
16201 TTTTGGAGA GTACAATCCA TATATTTTGC AGAATTTCCC TTAGTTTGGG
16251 TGTGTCTGGT TTTTCCCTAT AAGATTCATT TTATGCATTT CTGGCCAGAG
16301 TACCACAGAA GTACTGTATA TCTTACCAGA AAGCCTAAGT GGCATTTGCA
16351 TTTTCTAAAT GATCAATTTT AATATTATAT GGAAAGCAGA GTCAGAGATT
16401 CTCACATATG TCAAGATATT ATAAGTATTC CTGTTATATT TATTCTCCAA
16451 TTGCTTTTTT TCAAGAAAAT TTGTGGCCTT TCAGCTAGCT TTTCAAAGTG
16501 GAAGTTACTA CATAACATTA GGATGGGAGG GGTGGGGAAG AGCTTTATTA
16551 AAGCTTTAAG ATTGAGCTTT TGAGTATGTG TTGTATGTAA ATGAAAGTGG
16601 GCATTGATGC AGGGATTGGG CCTTTAAACC TTTGGCCAAG AATGGTATCA
16651 ATTATTATTA TTATTATTTT TTGGAGTACT TCTGCTAAAA CACTGAAATC
16701 AGTGTGCCAC TCTCCTTTTA GAAGTTTAC ACCTTTCCAA GGTACACTTT
16751 TTTTTTTGGA GACGAGTTTT GCTCTGTCGC CCAGGCTGGA GTGCATTGGC
16801 GCAATCACAG CCCACTTCAG CCTCTGTTTC CCAGACTCCA GCAGTCCTTC
16851 CACTTCAGCC TCCCAGTAG CTGGGATTAC AGGTGCACAC CACCATGCCC
16901 AGCTAGTTTT TGTAGAGATG GGGTTTTGCC CATGTTGCC AGGCTGGTCT
16951 CCAACTCCTG CGCTCAATCT ATCCGTCCTC CTCAGCCTGC CAAAGTACTG
17001 GGATTACAGG CGTGGGCCAC CACTCCCGGC TTCCAAGGCA GGCATTTAAA
17051 TGTAATAAAT AGGGAGATAA GCAAGAACCC TGTGGACCT GGTAGAAGCA
17101 AACATTTATT AGTACTATTA CGTTGTTTAA AATATTAGCG CCTTCTATAT
17151 TCATGTCCTC CCAGAATTAT CAAAAACCT ACTCTATAGT TTATTTGGCT
17201 TATATCTCAG GAGTAATAAA ATTAGTTAAT AGTATTGGCA TCGTGGTTCT
17251 TTGTGTATTC TCCCCTTATC CCACCCCAAG TTGATTTAC ATGATCTCTT
17301 GATCTAGTCT AAGAATGTTT ATAGTGATTA CGAGAAGTTC AGATTCTGGC
17351 TTTAACATAT ATAATTGTTT TTTAATCTGT AAACCAAAGA GAATGAGTTT
17401 GTTTAAACTA GAAAGATGGC AAGAGTAGTC TGGGAATTTT GTTCCATTCC
17451 TTAAAGTCC TATAATAAAA TAAACATATC TTGTGTTTAA TTTTACAAT
17501 TTTTTTAAAC ATTAGTACAG AGTGCCACTT CTTATATTCT ATATCAAATA
17551 ATGAGCTACA TTTTCAATAA TAACCTCTGA GTAATTTTGG GCATTAAAT
17601 GCTGCATTAC AAAATAATTT GAGGATATAA TTTATAATCA CTTATGCTAA
17651 AATCACCTAT TTGAAATTAT GTATGAGGTT TTCAAAGTTT ATAGTGCTTT
17701 GGAAAAAATT TAAATGTTTC TTTGTTTATG TATCTTTATT ATAAGCTGTA
17751 GCATATATCA TGTAGTTGTC AAGGATGCTG ATAGATACTT AATATTTAAA
17801 GGAGACTTGT CTAAGTTAG CTGTCCAGGA CTAGAATCTG GGCCTTTTGG
17851 TAACAGCTCA TTGCTCTATT TACTTAAATG ATGATTGGAT TCGTTAGAAT
17901 TTCTCTATTT TCATAGCTGT CTCTATGGTT CTATGAAAAT ACTGTGTGTG
17951 TGCTTATACA TATATGTATA CCTGTAAGTA CAAAGTAGAA AATGAAAGTT
18001 CATTTTCTGC TTTTGACAAT TGAATCCCC AGAGATAACC GTTATTAATA
18051 TGTTGTCTCA TGTTGGTCA TACTGTTTTC TCTGTATTCT GTGTATTACT
18101 GTATAAATTT TACACAGTAA TTTGCATATT AAAAATGCTG GTCTACACCT
18151 GGCCCTTTTT TAAAACTGC AATTTATTAT GGCCAATTTT TTATACCAGT
18201 ATATATTGAT CAACCTTATT CTTTTTAACT GCTGCATTTT ATTCAATTAC
18251 AATAGATGAG ACATTTCCAT TGGTTTGAAT TTTTCAGTAT TACAGATAAT
18301 GGTTCAATTA AATATTTAAG CTTTGTGCA CTTGTAGAAT TAATTCCTAG
18351 ACATAGAACC CTTATATTTT GATAGGTATT TCCAAATTTT TTCCCAAAAT
18401 GTTTGTATCT CTTTACTTCC ACTCTCAGGT CTAATAATTT TCACTTGGAT
18451 TATCATATTT CTTACCCAGC CTGTTTTTTA CACTCTAAAC TCTTTTTCTT
18501 TTCTTTTTTT TTTTGGAGA GCATCTTGCT CTTGGCCCGG TTGAAATGCA
18551 GTGGCACGAC GACCAACCTG GGCTCAAGCA ATTCTCTCAA CTTAGCCTAC
18601 TGAGTAGCTG GGAATACAGA CACATATCAC CATGCCCAGC ATTTTTTTTT
18651 TTTTTTTTTT GGATTTTATG TAGAGATGAG GTTTTGCCAT GTTGCCCAAG
18701 CTGGTCTCAA ATTCCTGAGC TCAAGCAATC CACCCATCTC AGCCTCCCAA
18751 AATGCTGGGA TTACAAGCGT GAGCCACTGC ACCTGGCCCA AAAGCTCTTT
18801 TTCTAATAGC AATATAAATT GTCTTTTACA GACTATACTC ATATATGTTT
18851 CTTCTTTCAG AAATAGGTGT TAAGTGATC TAACATGGAA TGTATAGCTA

FIGURE 3, page 6 of 27

```

18901 TAATTCATCAT TGTGAAACCA TAGCCTAATT TATTTTCATAT TACAATTTAA
18951 AATTCATATT TTTTAGGAAG TTTTCTTAGA TTAATCCGCC TAGTTCAGG
19001 TGCTACAGTC CCAAGATTTT TTTCTTTTAA ACAAATTAAA TATAGGTAAC
19051 ATGACTAGAA TTGTAGTCAA AGAATATTGG AACCTTGGA CTTCAGTATT
19101 TGAACCTTAT TTTGAAATAT AATTTGTTAT ATTATAAAAA TATTATAATA
19151 TATTGCACCT GGAAGTTAGG GGCAGTTTTT TTTAATTCTC TTTGTATCTG
19201 CTACACTGTA AAGTGCTATT TATGTAAAAA ATTCCTAATA GAAGTCTTCA
19251 GTTGTAAGT CTGCTGTACA GACTTTAGAT CAGGGATTGG CAAACTATGA
19301 GCCATGTGCC AAATCCTGCC CTTACCTGT TTTGTAAATA AAGTTTATC
19351 AGAACACATT CAGACTCATT CATGAACATA TTGTCTATGA TTTATTTTCT
19401 GCTACTATGG CAGAATTGAG TTGTTGCAAC TGTGTGGCAT CCAAAGCCTA
19451 AAATATTTAC TCTCCTGGCT CTTTGCCAAC CCGTTTTAGA TTATGAGCAC
19501 TTTGGCATT TATGTTTTT GTTTTCTTTC TATAGCACAC AGTAAGATGT
19551 TCTGCCACA TTGTGCATAA TTTATGGGTT TATTCAAGGA TTTATGCAAG
19601 TGTAGCTGCA AGAAAAAAC CTAGAAGTGA ACTTGCTAGG TTGAAGAGCA
19651 TCTGTGTATG TTAAATTTTG TTAGCTTTCG CCTTCCCAA GGGATTATTC
19701 CATTTTCATAC TTAAACTACT AATTTTGTGA TAGGACTTCT TTCTCCATAG
19751 CTTTGCTAAA TTAATGCATT CACACACTTC ATCTTTACTA ATCTGATAGA
19801 GGGAAATGAT ATTGTGGATT TGATTTGCAT TTCTTTTAT GTGTTAGCTT
19851 GAGCTTATTT TCATATTTAA AAGCCAATTG TATTTCTTTT TCTTGAGCTA
19901 TCTTTTAATG TCCTTCCTGA TACATTTCTG AAGTCTGTGA TACTCATATA
19951 AGATATATGG TGAACATGTG TCAAAGATTT ATTTGACTCT AATGAGGGAA
20001 CCCGCTGAT GACAAGGCTG ATTGAGAAGA GGATGTGTGA GATGAAGTGT
20051 ATATCATCAG TGAAGAAAG CAAATCTTA CAGGGCAAAA ACAAACCAC
20101 AACTCTAAGG GTTATTGTTT CTA CTGGACA GAATTCATT GCATTTTACC
20151 AGATAAAAAT TACTATTTTC AATTTATCTT TTACAAATCA TTTTCTAATT
20201 TTACAGAGTC TATTCCTTAA TCAGTAGTAA ATAGTCTTCA AAATCTCCG
20251 CAGCGTCAGG TGACTATTAT GCAGGCTAAT TGTTGACACT CGGGCTTGAC
20301 TTTAAGAGAA CATGCCATAA TCTTTTGGCC TTACTTCCAA GTTTTGATA
20351 ATTTTCTTAA ACACATTTT CTCTAATTC AATGATTTCA AGTGATATTA
20401 TTTCTTTTTT TTAAATTTT TTA CTATTTA TTGATCACTC TTGGGTGTTT
20451 CTCGGAGAGG GGGATTTGGC AGGGTCATAG GACAATAGTG GAGGAAGGT
20501 CAGCAGATAA ACATGTGAAC AAAGGTCTCT GGTTCCTTA GGCAGAGGAC
20551 CCTGCGGCCT TCCACAGTGT TTGTGTCCCT GGGTACTTGA GATTAGGGAG
20601 TGGTGATGAC TCTTAATGAG CATGCTGCCT TCAAGCATCT GTTTAACAAA
20651 GCACATCTTG CACCGCCCTT AATCCCTTAA ACCCTGAGTT GACATAGCAC
20701 ATGTTTCAGA GAGCAGGGGG TTGGGGGTAA GGTATGGAT TAACAGCATC
20751 CCAAGGCAGA AGAATTTTTC TTAGTACAGA ACAAATGGA GTCTCCTGTG
20801 TCTACTTCTT TCTACACAGA CACAGTAACA ATCTGATCTC TCTTTTCCCC
20851 ATATTTCCCC TTTTCTATTT GACAAACTG CCATCCTCAC CATGGCCCGT
20901 TCTCAATGAG CTGTTGGGTA CACCTCCAG ACAGGGTGGC GGCCAGGCAG
20951 AGGGGCTCCT CACTTCCAC ACTGGCGGC CGGGCGGAG CGCCCCCAC
21001 CTCCAGACG GGGCGGCTGC CGGGCGGGG CGCCCCCAC CTCCAGACT
21051 GGGTGGCCG GCGGAGACGC TCCTCACTTC CCAGATGGGG CGGCTGCCG
21101 GCGGAGGGG TCCTCACTTC TCAGATGGGG TCGCGGCTGG GCAGAGGTGC
21151 TCCTCACCTC CCAGACAGGG TGCGGGCTGG GCAGAGACGC TCCTCACCTC
21201 CCAGACGGGG CAGCCGGGCA GAGGCGCTCC TCACATCCCA GAGGGGGCGG
21251 CCGGCAGAG GCGCTCCCCA CGTCCCAGC GATGGGCGG CGGGCAGAGA
21301 CGTCTCTCAC TTCCTAGACG GGATGGCGGC GGGGAAGAG CGTCTCTCAC
21351 TTCCTAGATG GGATGGCGGC CGGGAAGAGG TGCTCCTCAC TTCCTAGACT
21401 GGGCGGCCG GCAGAGGGGG TTCTCACATC CCAGACGATG GGCAGTCAGG
21451 CAGAGACGCT CCTCACTTCC TAGTACAGGG TGGCGGCCG GCAGAGGCTG
21501 CAATCTCAGC ACTTCGGGAG GCCAAGGCAG GTGGCTGGGA GGTGGGGGTT
21551 GTAGCGAGCC GAGATCACGC CACTGCACTC CAGCCTGGGC AACATTGAGC
21601 ACTGAGTGAG CGAGACTCG TCTGCAATCC CGGCACCTCG GGAGGCCGAG
21651 GCGGCAGAT CACTCGAGGT CAGGAGCTGG AGACCAGCCC GGCCAACATG
21701 GCGAAACCCC GTCTCCACCA AAAAACACAA AAACCAGTCA GCGGTGGCGG
21751 CGGTGCCTG CAATCCCAG CACTCGGCAG GCTGAGGCAG GAGAATCAGG
21801 CAGGAAGGTT GCAGTGAGCC GAGATCGCGG CAGTACAGTC CAGCCTCGGC
21851 AACAGAGGGA GACCGTGGAA AGTGGGAGAC GGAGACGAGG GAGAGGGGGA
21901 GACCGTGGAA AGCGGAGGT GGAGACGAGG GAGAGGGAGA GGGATTATTT
21951 CTGTATGACT TAATAATGAA TTTCTAAGAG GTCACCTAGC TCACTGTTGT
22001 CTCTTCTAAA ACATACTCAT CTTTCTTTT CTCTTCTGTA GGAATCATT

```

FIGURE 3, page 7 of 27

```

22051 ATACAATGAC AAATGGAGGC AGCATTAACA GTTCTACACA TTTACTGGAT
22101 CTTTTGGATG AACCAATTCC AGGTGTTGGT ACATATGATG ATTTCCATAC
22151 TATTGATTGG GTGCGAGAAA AATGTAAAGA CAGAGAAAGG CATAGACGGG
22201 TAAGTGT TTT TAGTAAAAAT TTTTAAAAAC ATAGTGCATA ATTAGATCTT
22251 TTAATAATAT ATTTCTGCCA ATGATCTCAG GCTGCCAAAT GTTTACATTT
22301 AATATAAGTA AATGTCTACA TTTTCATATGT GGTACATGTT TTTTCTTTT
22351 TCTATGTTTA ATTTTTTTAG TTTACTTATA CCCTGTAAC TTTCCAGAAAG
22401 GATTTTCAGGT AGCTAAAAAA CAAAGAAATA CAATAAGAAG ACAAATAAAG
22451 AAGGAAAGGG AAAAATACAG CACAGGAGTT GGGGGGAAGA ACAAGCCAAG
22501 TTCCAGATAT GGAGGTCAGC ATGATTTTGG GCTTTGAGCA GCCCACCAGC
22551 TAAGGCAAAA AAGGAACTC ATTGCATAGC TCTTACCTAT GGAAAAAGAA
22601 GAAATCTACT GGGGGCAGAT GGTCTGTGG GATTTTGCTG TTTTCTTTTA
22651 TCTCCTTTCC CAGCATTTGA TTCTGAGATA TTTCTCAATT TGGCTCCCAA
22701 ATAAAGCTTA TTGAGTGTG TAATGGTTTA CTGTTTTTTT TAAAAATGGC
22751 TTTAACATAT AAAAGTACAA CTTATGGATC CTTTTTGT TTTTGTGTTA
22801 CTTACTGATA ATATAATCCA AAATACATTT TTTATTTTGT ATTTATTTAT
22851 TTATTTTTGA GACGGAGTCT CAGTCTCTG CCCATGCTGG AGTATAGTGG
22901 TGTGATATTG GCTCACTGCA CCCTCCGCCT CCTGGATTCA AGCGATGCTC
22951 CTGCCTCAGC CTCCTGAGTA GCTGAGACTA CAAACGTACG CCACCATGCC
23001 TGGCTAGTTT TTATACAAAA TACGTTTTTT AAAAAACAAT TTTTTTTTTG
23051 GAGGTCGGGG GACTGTCGCC CATTCGTGTG CCCAACTGG AGTGCAGTGG
23101 TGCAATCTTG TCTCACTGCA ACCTCTGCCT CCCAGTTTCA AGCGATTCTT
23151 GTACTCAGCC TCCTGAGTAG CTGGAATTAT AGGTGTGTGC CATCATGCCA
23201 AGCTAATTTT TGTATTTTGA GTAGAGATGA AGTTTCGCCA TGTGGCGAG
23251 GCTAGTCTCA GACTCCTGGC CTCAAGTGAT TGGCTGACCT CAGCCTCCCA
23301 AAGTAGAAAA TCTTCTTGAA AAATAAAATT CCAAATCTCA AAAGGCCCTA
23351 TATAATTTTG GTGTTGGAAA TTTACTTGTC AATGAAAATG ACTATTTACA
23401 CAAATTATAA GCTTCCATAT TAATATATAT GTGTGTGAAC CTGAAATTCA
23451 AATTTTATTA TATTGTTTAT GAAAGGTACA GCCTCTGAGA TTCATCAGAT
23501 GGTATTTACC TTTAGGGCAT ATCTAAAAAT AAAATACAGT ACATGAAATC
23551 CAGTGCTTTA ATCCAGTGAT TCTTAACTT TTTGCTCTCA GATCCCCTTT
23601 AAATCTTAA AAGATATTGA AGAGCTCAA GGAGGCTTGG TTTACGTGGT
23651 TTTTATCAAT GGATATTTAC CATATTAGAC ACTGAACTG AGGATTTTAA
23701 AAAAAAATAA TTCATTTAAA AATAACAGTA ACAAACCCA TTACATGTTG
23751 ACATAAATAA CATTTTACG AAATATATT TTCAAAATTT AGTGAGAGAA
23801 TGACATTGTG CTACATTTGT TATAAATCTC ATTATTGTCT GGCTTAATAA
23851 AACACTGCTG GATTCTCATA TCTGCTTTG NNNNNNNNNN NNNNNNNNNN
23901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24851 NNNNNNNNNN NNNNNNNAAA TATTGATTCA CTGATTATG TGGATCTTTT
24901 AAATGTTGAC ACTATATAA TATAATACAA TATTTTAAAA ATCACATTTG
24951 TTAATTTTAC CTTTGATCTA TTCAGAAAAG ACTCTAAGTA TTGGGAACCT
25001 ATCATCCTCA CAGTGATAGA TACAAGTTC CTAAATTTCT GATTTTACT
25051 GGAGAGCTCA AATTCTATCA TTGGAACAAA ATACACATTT ATTTAACTTA
25101 AAAATGACAG GATTCTTGG TTTTATTATT GAGAAAATAC CTGTCAAATT
25151 CCCAAGTCTG GAAAACCATG GTTTGATGTC ACTCTTTCOA GTAAAAATGG

```

FIGURE 3, page 8 of 27

25201 CATTCCATGT AAGAAGTGTC TAGTTTATTA TGCAACTCAA ATAAATTACG
25251 CAAGTGCTTT TCTTTAGGAC ATAACCTCAT ACATACTTCC ACAAGCAGCA
25301 GATGTGTGTA GTTATGCATA GTTCCTTATG CATGGTTCTT ATTTTCATCAC
25351 AAAAAATATT AAAAAGACTC AGTGATTGAG ACGTAGCAGT TTTTACTGCT
25401 TCATCAAAGA TGCTCTTATT TGAAACTGGC ATAATATGAT TTATTTATTT
25451 GATTTTACTG GGAAGCATGG CAGTCAAGAA TGTAATGACT GCCAGTACAT
25501 TTGAGTGCCA CTGCTTGATT TTTGCTATGG AGTCAGCAAT TTTGCCACTG
25551 GTTTTGCATT TTCAGTAAAA ATGTCAACAC AGTGAAAAAG GCACATAATG
25601 TCTTGTATTA TTTTGTAAAC AGTTTATCT TGCAGACCCC TTGAAAAGGT
25651 CTCGGGGATC ATCCAAGGTG CCAGTAGACC GTACTTTGAA AATCACTATT
25701 TTAATCCAAA GTGCCTAGAT CAGACACACT ATAAATCCTG TGTCTTGTAT
25751 GATCATTAGG TAAATACATT TGTACTTAGA AGTATACATT CAGAGACATT
25801 AACAGTATTC AGGTTGGGAT TTAAGTATAT TTAAAGTGT GGTACCTAGA
25851 GAGTATCCAT GACACTATGT TCATAAAATT TTAGAGAAAA CTGAGATCAA
25901 AGGAAACCAA AACAGGCTGG TCATAGTGGC TCATGCCTGT AATCCCAGTG
25951 CTTTGAAGG TTGAGGCAGA GGATCGCTGG ATCCCAGGAG TTTGAGACCA
26001 GCCTGGGCAA ATATGGAGAC TATCTCTACA CAACAAAACA AAAATTAGCT
26051 GGGTATAGTG TCTTGCGCCT ATAGTCCTAG CTACTCGGAA AGCTGAGGTG
26101 GGAGGATCCC TTGAGCCTGG AAGTTCTAAG TTACAGTGAA TTATGATTGC
26151 ACCACTGCCC TCCAACCTGG GTGAAACAGC AAGACCCTGT CACCCTCCAA
26201 AACAAACAAA AAACACTTTT TTCTCTGAGT ATGTAAATGG TTAGTGTACA
26251 GTCCTTGAAA ACATTTGCAA TAGTATAGCA ATATATGAAG TAGCCAGTAT
26301 GTGCTCTAGC TAATTTTATC AATCATCTCT TCCTAGACCA ATCAAATATT
26351 TTTCAATATT TTGATCCATG CTTATATGAA CAAGATTTTT TAAAGCTGGA
26401 AAATTCCACA CATTTATATA CTTACTATTG TTCTTAAAT TAATTTTTTT
26451 TTTTTTTTTT TAAGCAGAGT CTTGCTCTTT TGCCAGGCT GAAGTTCAGT
26501 GGGGCGATCT CGACTCCCTG CAACCTCTGC CTTCCAGGCT CAAGCAGCTC
26551 TCGTGCTTCA GCACCCCAAG TAACTGGGAT TACAGGCATA CGCCACCACA
26601 CTGGCTAATT TTTGTAGTTT AAGTAGAGAT GTGGTTTCGC CATGTTGGCC
26651 AGGCTGGTCT CAAACTCCCG GCCTCAAGTG ATCCACCTGT CTCAGCCTCC
26701 CAAAATGTTG GGATTACAGG TGGGAGCCAC TGCGCCCGGC CTACATTAAA
26751 TTTTAAAGCC TTTCTATGTC AGTGCAATATA CCCAACCTAA TTCTTTTTTT
26801 CCGTGAACCT TTTTGTATG CTTGTAGCCT TCCTACCCCA GATTATTTTCG
26851 AAGCAAATTG TCATTCTGTA ATTTCAAATA TTACTATTTT AGTATTTTAC
26901 AAAATGGTTG CAGTTTAATT GTTGTTCCTT TTTTATTTAT TAGCTTGCAT
26951 ATTTCTATAG AGAGTTTACC CCACATCAAC CATTTGGATT ACCTGAAGTA
27001 AGGGTGGTAC AGGAAAGGGA GAAATCTTGA AATACTAGGT TCCTTAGCAT
27051 CCTCAAAGTT GACCAATGAG ATTTTTTGCT TGTGTTGGTT TTTTTTCTG
27101 TGTCTTCTGG ACTCATGGAT TTAAGTATAT TTGTGTTTAT ATCATCACTG
27151 TTATATTCTT TATTGATGTT CATGTTATTT TAGATTAGTG GGAGCTTTTT
27201 TAGTTTGCTA TCTGTGTCCT TCGTCATGTC CTTAGATAAT CCTAATCCTA
27251 ATCTGATTG ATCGTAGACA TTTCCCGCAG CAAACCTGGA ATCAGCCATT
27301 TCTCAAGGAG CTCTCTGATT CCATTGAAGG AAAATATAAT ATAGGTACAA
27351 TCTAGGCACT AGGTGATACT TGTTACTTCT GGGTTGGCTA TTGTTTCTAG
27401 CCTCCTAAGT TTATATGACT GTACTAATTT GAATTCATAA CTATGGGACT
27451 AAACCTCTAA TTCTTAAATC TGCATTTTCT TTAAGTCATG CCAAAAATCT
27501 GAACATCACA AACATAGTCA TTTGCTTTAC CCCACAATAC ACACATACAA
27551 CATTGTCAGT ATAACAGTAC CAACACCATC TCCAACAATA TGCCTACTGA
27601 AAAATTTTAG GTAATCTGTC TCCAGCCTCC CAGGTAGCTG GGAAGTGCAGG
27651 TGCACACCAC CATGCCTGGC TAATTTTTTT TTTTTTTTTT TTTTTTAAGA
27701 GACTGGGTCC TTGCTATGTT ACTCAGGCTG GTCTGAAATT TCTGGCCTCT
27751 AACAGTCCTC CTGCCTTTGC CTTCCAAAGT GCAGAGATTA CAGACCTGAG
27801 CCACCAAGTC TGGCCTATCC TTTATTTATT CCACCAAAGT TATTTATACA
27851 AATTACTTTG TTGTAAAGTC CCTTGGAATA GTTCTTCTG TGGCATTATG
27901 TTACCAGTTA GATGCACCTT TGATTCATTT AACTTTACTT CAATTTTAA
27951 GGTTTGCTTT TTAGATTTAG TTTTGTTTTA TTATACATAT ATGAAGTATT
28001 TCCACGGTTC CAAAGTTAAA TGAACAAAAC AGGCATGTTT AAAGAAGTCT
28051 AGTTTCTATC TCTGTCCCAT CCAACCCATT GTCTTCTTCC CCTTATAAGT
28101 AATAATTTAC ATTTTAACT TGTGGTTTAT CTTCTGATTT TTAATAATAT
28151 AAGCATAAAT ATTTATATTC CTGTCTTTTA GCATGCTTTT AGCCATCTTG
28201 CTTTTTTCCT GTATAATGCT AAATATATCT CATTCTTTTT AATTGCTGCA
28251 GAATTTCTCA TTACATAGGT ATACTGCAAT TTATTTATCT GATGCTATGT
28301 TGATGAACAT TTAAATGATT TCCAGATTTT AGGAACGGTG ATGATTGAAC

FIGURE 3, page 9 of 27

28351 TCTCTGTACA TATATCTTTT TTACTTGGTA CACTCCATCA AGCAACTACT
28401 TAAGTGACTG ACTATGATGC TGTGCAAGCA GTTATATAAA GAAAACAGCA
28451 GTGACTCAGC CTGAAAACGG CTTAATATTA TCATGTTTTC TTACACATTA
28501 TTTTATTGA GGAAGCAA CATGGAGTTT AGTGATTATT TTTGAAAGAA
28551 ATAACCTATT TCTAATTCTA AAGAATGGTT ANNNNNNNNN NNNNNNNNNN
28601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
29001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
29051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
29101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGAGTCTAG
29151 CTCTGTCACC CAGGCTGGAG TGCAGTGGCA CGATCTCTGC TCACTGCCAC
29201 CTCCGCCTCC CGGTTTCAAG TGATTCTCCT GCCTCAGCTT CCCAAGTAGC
29251 TGGGATTACA GGCCTTCGCC ACCACACCCA GCTAATTTCT GTATTTTGTAG
29301 TAGAGAAGGG GTTTCACGTG GTTGGCCAGA CTGGTCTTGA ACTTCTGACC
29351 TCGTGATCCA CCTGCTTCGG ACTCCCAAAG TGCTGGGATT ACAAGCGTGA
29401 GCCACCACAC CTGGCCAAAAT ATATGGGTTT CTAAAGCAAC AGTCTAGTA
29451 CAACAGAAGA GAGGTGTGA CTAGTTAGGG ATTTAGGTTT AGAAGTACAT
29501 TCTTAGTAAG AGAGGTGAGA CTTACCTTCT TGTGTTTTAG TATAGTGAGA
29551 TCTGGATCAA ATCTATTACT CTTATTAATC TCCTAACTTC CTACACTATA
29601 TCCAGTAGAG GACACTTTTG CCTTACACAG TAAAGAAAGA GCCTCTGGAC
29651 TCTACCAATG GGATCGGAGC TCTCCAAACC TGCATATTAA AAGGCCTATA
29701 AGTTTTGGGG GGTCCCTTTG TCCACATGAT TATTCTGTAA TACATTGTAT
29751 TTATGGACAT GGTATTATTA TACACAGATC CTGTCTTTTA AAGAACATTA
29801 TAATCCACTT AACTGCTAGG ACCAGAGAAT GACCGATAAT TCAAACCATA
29851 TTGTCTTACA GAAGACATAT ATAAAAGATG GTTATGTGTA CCAATTGAGG
29901 TTCAAATTTG ATTCAATTTA AAACAATCTA GGCCAGATTT TATATAGTTT
29951 GTGGACCTTT TGCACTCAAA TCTCAAGGTT CTTATTAAAA TGCAATCTTT
30001 GGCTGGGCAC GGTGGCTCAC ACCTGTAATC CCAGCACTTT GGGAGCCCAA
30051 GGCAGGTAGA TCATTTGAGC TCAGAAGTTC AAGACCAGTC TGGCCAACAT
30101 AGCGAGGCCC AGTCTCATTG AAAGAAAAAA AATTTTTTAA TAAAAAATAA
30151 AAGCAGATCT TGGGTAAAGA CATGTAGTCT GGTTTACAGG TATTAACAAC
30201 TGTCTGTAAT GTAGTGATTT TGCTCCAGAC TTACCTTTTC CATTATTTAG
30251 TTCTGAAATT ACTGTTCTAT GTATGGTAAA TGAGAAAAAT TGCTAGATTC
30301 TAGAACTGTG GCTTCTATTC ATAGTTGGAA AAATGAAGCA TAAACATTTT
30351 TAATTCAGA TCAACAGCAA AAAGAAAGAA TCAGCATGGG AAATGACAAA
30401 AAGTTTGTAT GATGCGTGGT CAGGATGGCT AGTAGTAACA CTAACAGGAT
30451 TGGCATCAGG TAAAGAAAAA TTTTCAAGCA ATCCTTTTTT AGTTAACAGA
30501 AGTATAAACT GTTCTTCCCT CCTTCCCTCA ATTTTTTTTC AGGTACCATT
30551 GGATTTTAAA AAGCATTGTG TTCTCTTCTT CAAAAATCT CCTTAAATAT
30601 AAGACTAGGA GGCAGAGGCT TCCAAGTCTA GTCTTGGCTC TATCACTTTA
30651 CGTGTATTATC CAGCTTGGTT GATCTTTCTG GACTCAGTTT CTATATCTGT
30701 AAAATAAGTG GTTTGGATCA GATGATCAAT AAAGTATCTT TTGATATTAA
30751 CATCGTAATA AATAGCTAAT ATTTCTTGAG TGCTTCCTAT GTNNNNNNNN
30801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNTGGAAGAT
30851 TATGTTTACA AGACCATAAA AATTAATAAT TTTGTGGAGA ATAAAGTACT
30901 GATAATTCTA ATTGGCATGC ATAGTAATTT TATGGCCTCT GTGTATGTAA
30951 CCCACTGATC TCTTTATGTA AGAAGGACCC AGATTTGACC ATAAATTTGT
31001 GTATTTTTTA TATTCTCACA ATAAAATAAT CTTGATATAT GGTTTTCTGT
31051 AATTTAAGAA AATATTATTC CTATGAGTTT CAATAATTAT TTCTAATGGA
31101 CATTAATAAT TAATGAAATT GACATCATTT ATAAGTCTGT TAATTAAGTT
31151 ATCGATTGAA AATTAGATTT GTGAACCTCC TGCCAAGTAG CTGTCTTTTG
31201 AAGATATTTT AGTATCTTTT AAACATTGTT TTTTCAAGCA CAATTAATTT
31251 GAATGATGTA ACTTTTAAAA ATTCCAAACA AAAATAGCAC TTTTATTGTA
31301 AAAAATAACT CTTTACAGTT TATAACTAAA ATTTGAAAAT CTTAAATTTA
31351 TATGTAGTTC ATAAATGACC CTTTATTTAG GAGTCTCCTG CTTTCTACTT
31401 GCGTTTTAAC TAGATTGTTT TCGACTCCCA AAAAATTGAC TTAATTTTTT
31451 TACCATCTCC AACATGTTTT TATAGGGGCA CTGGCCGGAT TAATAGACAT

FIGURE 3, page 10 of 27

31501 TGCTGCCGAT TGGATGACTG ACCTAAAGGA GGGCATTTCG CTTAGTGCCT
31551 TGTGGTACAA CCACGAACAG TGCTGTTGGG GATCTAATGA AACAAACATTT
31601 GAAGAGAGGG ATAAATGTCC ACAGTGGAAA ACATGGGCAG AATTAATCAT
31651 AGGTCAAGCA GAGGTAAGTC TTGCTTTGTC TCAAGATGAA TTAATAATTG
31701 ATATAGCAAA ATGTTTCCAA TTCATTTAAT TATAGAACTA ATCACATATT
31751 AGATGATTAC ATACACATCA AATGGATCCA CCTCAACAC ATTGCAGCAA
31801 GAAAGAATTA AGTGCAATAT TGTTCAGT AGCTTTTTTA TTAGTTAACT
31851 GCATAGTCAT ATAACAAATC CTCTGGATTG TGGTGCAAAT ATATTTGAGC
31901 TGTAGTAGAA AAGAAGTGAT AGTTATTGCA GTAAGATCTG TGTAAAGTTA
31951 CTAAGAAGTC AAGTTATTAA AACTAATATA TTAAGATAGT TGGGAAGTTT
32001 GAATTATGAA AGTATTATCA AATAATTTAG TAAATCAAC CTACGTAGAG
32051 ATACATTGAA GATAATCAGA CATTTTATT TGTGGCATT CAGCATTTAA
32101 ATGATTGATT TACTATGATC TACAAAGAAC ATTTTAGAAC TTAGGATGTT
32151 ACATGTATAT TTTTACATG ATGACATGGA TATATTTTTT AAATTTTGT
32201 TTAGCTGAAC TTTAGAGCTA AAAGGTATAC ATTTGCGGTA AGATGAGTAG
32251 TATGCTGTTT CTCACCTGGC TTAATTGAAT TGAGTTTAAT GATCTGGAAA
32301 GTTGACGAGC AATGAAATCT GAGTGGTGAT GCAATTTGTT TCCACTGTTT
32351 CCAAAAAGTG GTTTGTAGGC AGAGATTGAA GTATAGCTGA GATGTGTTGG
32401 TAACAAGACT TTAGGGATTA GGAAAAGAT TAAATGTGCT CAGGGTTCCT
32451 TGGTATATGT AGGCATTAAT TTTTGGACTC TACTTAAATA TTTTGTTCAT
32501 ATAAAGTTT TATTATTGTG GAAATAAACC AGGAGACTTT TACACATTTT
32551 ACTGAAGTTT CTTTCTTTC TTTTCTTTT TTTTCTTTT TGGCCGGTGG
32601 GATGGAGTCT CACTCTGTTG CCCAGGCTGG AGCGCAGTGG CACGATCTCG
32651 GCTCCCTGCA ACCTCCGCTT CTGGGGTTTA AGCGATTCTT CTACCTCAGC
32701 CTCCCGAGTA GCTGGTATTA CAGGCGTGG CCACCATGCC CAGCTAATTT
32751 TTGTATTTTT AATAGCAACG GGGTTTCACC ACATTGGCCA AGCTAGTCTC
32801 GAACTCCTGA CCTCAGGTGA TCCACCCGCC TCAACCTCCC CAGTGCTGGG
32851 ATTACAGGCG TGAGCCACCA TGCCTGGCCG TTTACTGAAG TTTCTTATGA
32901 CAAGCATTTG CATTAGAGGT GCAATGTAAA TTAATTCAT ACTCTCGAAC
32951 TATTTTCTTT TTAGGGTCTT GGTCTTATA TCATGAACTA CATAATGTAC
33001 ATCTTCTGGG CCTTGAGTTT TGCCTTCTT GCAGTTTCCC TGGTAAAGGT
33051 ATTTGCTCCA TATGCCTGTG GCTCTGGAAT TCCAGAGGTA AGCCAAGTAA
33101 TATTTAGTGT CATTAAACAT TATTATGATG CTTATCTTTT TGACCTTAGT
33151 GATAATAAAA GTTGGCTTTT CTGGAGGGAG GGGATAGTTT GTTCATAATA
33201 TGAAAAAAAT ATTTTTTTAA GTATAAGCTG ATGGTAGACA TCATTGAAAA
33251 ATATTGTTCC CCATAGTCAT TTGGTCATTT ACTGTGAAGG CTGATTTTTT
33301 TTTTCTCTCA CCACTAATTT AACACATGAC TAGGCAAATT TTCAGACTAT
33351 TTAGTTAAAC ATCAAGAGCC TGAAGAAGT ATCTTGTGAC CTAATGTTCT
33401 TTGACGGGTT AGTTGTTACT TTGCTGTTAT GACCTGAAT TTTTTTTTTT
33451 TGAGACTGAG TCTTGTGCTG TCGCCCAGAC TGGAGTGCAG TGGCGCAATC
33501 TCAGCTCACT GCAACCTCTG CGTCCCAGGC TCAAGCAATT CTTGTGTCTC
33551 AGCCTCCTGA GGAGTTGCGA TTGCAGGCAC CTGTCAACAT GCCCTGCTAA
33601 TTTTGTGATT TTTTGTGTTG TTTTGTGTTT TTAGTAGAGA TGGGGTTTCA
33651 CCATGTTGGC CAGGCTGGTC TCAAACCTCT AACCTCAAGT GATCACCCGC
33701 CTCAGCCTCC CAAAGTGCTG GGATTACAGG TGTGAGCCAC CACAGTGGC
33751 TATGACCCTG ATTTTGATTC ATTCACTTT TATAATTACC TTTTGATTAG
33801 ATAAAGTTAAT TATCTTGAA TTTGGCCATT TTATGCTTTG AGAAAGTAGT
33851 TAATCACAGT GGGTCAACAG TACAACTTT TGGGTTTAT TTTTCATCAC
33901 AATAAAGTAG AGTTATACAT AGGATTGATT GAACCTGATT TGAACCTATC
33951 TCTTCTCTTT TATTTTCTG GAGTTAAATA AGTTACCAAC TTTTCTCTAA
34001 TACATTTCTT TTTAAATGG AATTGTATTG ATCCTTTAAG TTTGTATTAA
34051 GAATATCTTT CATAAAAAGC AATATCATGC AGTATATAAC AGTTGTTACT
34101 CATTCTTGAT ACATAAAAA CTATTGCACA TAATTACAGG ACCTCAGAGA
34151 AAACATAATA TTCTTATTT TAACATAATG GCCAAAATAT ATTTAAAATA
34201 TTATGCTTAT TTTTACAACA GAAATATTCA AATTTGCCCT TTTTGTGGGT
34251 ATGTAATTAT AATCCTTATA ATTAAGGTCT GTATTCAATT TAACATGGCC
34301 TGATATTTTG ATTTTGGCCT GAGATAGTGT TGCCCTCTCT CTTTCTTGG
34351 GTAGAGAATT AGATTATAAT ATCAATTTAT TATATGTAGC ATAATAGGCA
34401 AGTTTTTCGAA AAATTAACCTG TAAATTTTTC TGTAAGTGC TAAATTTGTC
34451 AAGGTTGTTT TTGTGCATAA AACAAGAAAA TAACTGGAT TCGTTACATT
34501 CTCATGTTTC TTAAGGACA TTAAGCTGCC TTAATCTTTG CCTGTAGAT
34551 TAAAACCTATT TTAAGTGGAT TCATCATCAG AGGTTACTTG GGAAAATGGA
34601 CTTTAAATGAT TAAAACCATC ACATTAGTCC TGGCTGTGGC ATCAGGTTTG

FIGURE 3, page 11 of 27

34651 AGTTTAGGAA AAGAAGGTCC CCTGGTACAT GTTGCCTGTT GCTGCGGAAA
34701 TATCTTTTCC TACCTCTTTC CAAAGTATAG CACAAACGAA GCTAAAAAAA
34751 GGGAGGTAAG TGTCTTTTGT AGTTAATTTG ACTGAAAAAT ATATATTATA
34801 TAGTATTAT TTAAGTAAAG AATTTCTTAG TGTAATAATA ATAAATCTG
34851 TATTCAGATA AAAAATTTTG AGATTGTGC TTCTGTTTTT CCTGAATAAT
34901 CTATAACATC TTTCTAGAAT CCATTTCCAG TGCTGCTCAG TTCGTCTTAC
34951 ATTTTAGAGA AGCTTTAGAT AGACAGCTGG TGTCCATTGG GTTTCAGCTG
35001 CATTTACGA AGATCTTCCT GTTATCACTT TACCTTACAT CTTTCCTCTT
35051 CTGAAGTGTT TTCTAAGCTT AGCTTTGTTT TTTACTCTTA CTTTCAACAT
35101 TAAGAGGTTG GGAAATCTTA ATAGCTATGT TTTCTCCTG GAGGCAGTGT
35151 CTGGTGCCAG TGTAAGTGGT GTGTGATATG AAAAATGCTA TCCAGTGCTA
35201 TGGGGAAGTT CTGAGGGCCT TTAGAAGCTC TTGAAGTTTA AATCAGAAAT
35251 TCACATTAAA GAGATTACAG GAAATCCTTT TCATTTGATT GTTTAAGGCA
35301 ATTTCCCTTA CCATTTCTTT AGGCCAGCCT GAGATCTTCT ACAAGACCTT
35351 GAAACCTTAT ATATATTATG GATTTCTCTT GATGTTTCCA TATTGCTCTG
35401 GGCATTTTCC TGAATCCTTT ATATTAGCTC TAGACTTTGG GAGCCCAGTC
35451 CTTTCTTATT TTCCAAATCT AAATCTACAG CCCTAGATGG TACAGAGATC
35501 TTTGAGTTTT TAAGATATGA TTTTTTGAAA AACATCTCAT TAAATACTGG
35551 CAGAACCTTT TCATCTTGTT GAGTTTTTTA ATGTACTGTA ACCAAAAAAG
35601 TAGAATATTT TATCAAACTG TTTAATCTTC AATTGAAATA ATTCTAGTAC
35651 ATTTTAATGT TCGCATTAAG ATATTGTCTT TGCATTGGAC GTAGATATCC
35701 CAAAAGTGGG AACTTTCAGA TTGTCGTAGT TTCATCTCTG AATAATTGTG
35751 ATTCCAGTAC TTTATAACAA AAATAGCTAG CATTATTGAT TACTTTCTGT
35801 GTATCTGGTA CTGTGGCAGA TACTTTACTT GGATTTTAAT ACTTAATTTT
35851 ACAGTAATTT AGTAATATGG CCCTGTTATC CTCATTTAGT GATTAGTAAA
35901 CTAGGGCTGA AAACAGCTAA CTAACCTGCC CGAGACTACA TACCTAGTAA
35951 GTGGTGAAC GTAGGTTAAA ATTCATTTTT CTTTGACTTC AAAGTCTGTG
36001 GTCTTACCTA CTTACATTAC TGCCCTTACG ACTATGTGGG TATATATTTG
36051 TGTGTGTTCA AAACAACTC AAAACCATCC TGTAGCGTAG CAAGTTAGTG
36101 GCTAAGATGA AGCTAGAGCA TTTGCCTCCT CAATTCATTT CCATTACTTT
36151 CTGTTGTACC TTTATATTTT TTGGTAAAG TTTTACTTAT TCTAAGTTCA
36201 AAAAATGTAA TTTATTAGAT GTTTGAGAAA TTAAGTTTAC CTAAATTTTA
36251 ATGTTCATAC TGTAGTGATT AGTTAATGTT TAATACGTTG TTATTCTGTC
36301 ACCTTAGTGT ATATATAAAT GGCAAGAATT CACGGTAGT TGAAAGCATT
36351 AAGGTCCCAT AGTTTTGTGT AGACAAGAGG GGAGAGCGTT GATATTTTAA
36401 AATTAAATGC TTCTTAGATA CGTATGAAAT GGATTAAGAC ATGTATATGA
36451 GTTATAGATA CCTAGGTGTT AGTTTGTTG TAAATTCAGG ATCAGGACAT
36501 TCAAATAAAT ATGTTTGCTT TCCTCTTAGT GGAGGAAAAA AAAAGAAGC
36551 TAAATTTGCT CCCTTTCTCT CCCAAATAAG CAGAGTCTAC ATTTTAATGC
36601 CAACAATTTG ATTAAAACAA ATATTTATTT ATTTTAAATT CACCAAACCT
36651 TTATAAAGTA TTTACTGGTG CCAGGCACTG TTCTAAAGCA CTCTGTATAT
36701 ATTTACTCAG TCCTTAAGAG CTAAGTAATA TTATCAGTT TCCATTTTAG
36751 AGAAAAGTGA GGCACATATA GGTAGGTTA TCTACCCATA GCCATACAGC
36801 TAGTAAGTAG CAGAGCCATG ATTTCAACAC AGCAGCCTGA CTATGGAGTT
36851 CATGATCTTA ACCATTTACA GCTTAATTTT TATTATTTAT AATTTCTCTT
36901 CTGGAAATGT AACAATTGAC CATTTGAAGA AATACTTTAG GTAGCTTTGG
36951 ATATTTGCTG TATTAAAGTA GTGAAAGTAA TACAGACACT TGGCTGGGCG
37001 CGGTGGCTCA CGCCTATAAT CCCAGCATTT TGGTAGGTTG AGGCAGGCAG
37051 ATCACCTAAG GTCAGGAATT CGAGACCAGT GTTGCCAACA TGGTGAAACC
37101 CCGTCTCTAC TAAAAATACA AAAATTAGCC GGGCGTGGTG GCAGGCGCCT
37151 GTAATCCCCA GCTACTCGGG AGGCTGAGGC AGGAGAATCA CTTGAACCCA
37201 GGAGGTGGAG GTTGCACTGA GCTGAGACGA CGCCATTGCA CTCCAGCCTG
37251 AGAAACAAGA GAGAACTCT GTCTCAAAAA AAATAAAGGA ATACAGACTC
37301 TTAGAAAAAT AATTACAAAT AAAACCTTAG TGAAATTATA GGTATAGTTA
37351 GGTATAGTTG GCTTACAGGT GGGAAAGTAG CCATTACCAA CTGATAGACT
37401 GGGGAGCTGG AGAGAGGACA CGGAAGAGTG TCCTTGGATT TTTNNNNNN
37451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NAAAATTGTC TATATTCATT
37701 GCCTCTCCTT CTTTACACCC TATTCACATT AGTATATCTG GCAAAAATTT
37751 TTTTAACTG AATGGTAAAT GCATGACTGA CCTTTCATTT AAAGCCAGGA

FIGURE 3, page 12 of 27

37801 GAAAGAAACA AATCTTAATA GAAGAAATGA ATAGTTACCC TTTGCTTAGG
37851 GAGCAAGGAA ACATGCAAGT TAAATTCAGA AAATCCATTT GGAAAAATTCA
37901 AGTAACATGA AGAATTTTTA TTTGGTATGT TTGAATTTCT ATGAAATTAT
37951 GAAATAAGCC ATATCCTCTT TCTAGGTGCT ATCAGCTGCC TCAGCTGCAG
38001 GGGTTTTCTGT AGCTTTTGGT GCACCAATTG GAGGAGTTCT TTTTAGCCTG
38051 GAAGAGGTAG GTGAAAAGAA TACAACAATT AAAATTATAT ATAATTACCA
38101 TTACAAATAT ATTTACACACA TTTCAGTTTT GTAGGTGATG TAATAGGTAG
38151 AGACTTTGTT TTCAAATTTA TTTTCTAAA GTTGTTTTCC ACTCATTCTT
38201 AATAAAAAGT AAATGTTATT CATGCTCCAT ACCTGGAGGA AACTTTTTAA
38251 AAATTTATTA ATGTATGAAT GTTAGTAATT ATTTAAATC TAACTTTGTT
38301 GACATATTTA AAAGTAAGAA GATGTGAATT TGACTTAATA GAGGACATGT
38351 GAAACAATCT ATTTCCATTG GCTAAATCT GTATTTTGTAG TAGAGATGGA
38401 ATTTACCATT GTTGGCCAGG CTGTTTTTTT GTGGGGTTTT TTTGTTTTGT
38451 TTTGTTTTGT TTTGTTTTTT GAGACGGAGT TTCACTCTTG TTGCCCAGGC
38501 TGGAGTGCAA TGGCGCGATC TTGGCTCACT GCAACCTCCG CCTCCAGGT
38551 TCAAGTGATT CTCCTGCCTC AGCCTCCAA GTAGTTTTTG TTTAAAAAT
38601 TTTAATCAAT TCCTATGTTG AGTTTAAAG TTTTCCCAT GTGATTATTT
38651 CTGATACAGT TAGTGATGTT AAAGAAAATA ATTTTAGTGA CTTCAGTGGA
38701 TTATTTTGTT TTTGTTTTCT TAATAGGTGT TTAAGACTTT TCTTTTTACA
38751 TAAAAATGTA ACCAGGAATT TTTTTTTAAT TTTTTTTGAC AAATAATAAT
38801 TGTTTTTGTT TATGGGGTAT AATGTGATGT GTCTATACAT GTATACATTG
38851 CGGAATAATC AAATCAGAGT GATTAGCAA TCCCTCAAAT ATTTATTATG
38901 TCCTTGTTGGT GGTGAGAACA TTTAAATCC TCTTTTAGCT ATTTTGAAAT
38951 ATATAATACA TATTATTAAC TGTGGTCATC TTACTGTGCA ATAGAACACC
39001 AGAACTTAT CCTCCTCTGT AAGTTCATAC CCGTTGACTA ATGTCTCCCC
39051 TTTCCCTGTT CACCTCCCA ACCCCTAGCC TCTGGTAACC CCTATTCTAC
39101 TCTCTACTTC TATGAATTTA ACTCTTTAG TTCAAGATGT TTTTAAATGT
39151 ACTTTTTTCT TTTAGTTGTT TGTATTCTTT TTTTTTTTTT AATGTAGAAG
39201 AGGCAAATTA AATGCATTAT AAGTTAACAG GAGTTGGTGA TGGTACATTT
39251 ATTTTTAACT ACCATGATTG AATTGAATGT GAAACTCATT TTGAATATAA
39301 AACAGCACTA GGTATTCTAT TAGTATTTAT TAGACATTTA TGATCAATTG
39351 ATACTGTCAA TTTGTAATGA TGATCACCAT CTCGAAAAAT AATAATAACA
39401 TCAATTTTTT TTATTACAGT AAAATCCATT ACATGTAAAT TCTAACTACA
39451 GCAAAATTTA GAGCTAGGAT ATTTACCATT CAAGTTATAA TATATCAGAA
39501 ACATCTTATA AAATTATAGC ATTAATTTTT CTTTTCTTTT TCTTTTTTTT
39551 AGGTTAGCTA TTATTTTCCT CTCAAAACCT TATGGAGATC ATTTTTTGCT
39601 GCTTTAGTGG CTGCATTTGT TTTGAGGTCC ATCAATCCAT TTGGTAACAG
39651 CCGTCTGGTC CTTTTTTATG TGGAGTATCA TACACCATGG TACCTTTTTG
39701 AACTGTTTCC TTTTATTCTT CTAGGGGTAT TTGGAGGGCT TTGGGGAGCC
39751 TTTTTCATTA GGGCAAATAT TGCCTGGTGT CGTCGACGCA AGTCCACGAA
39801 ATTTGGAAAG TATCCCGTTC TGGAAAGTCAT TATTGTTGCA GCCATTACTG
39851 CTGTGATAGC CTTCCCTAAT CCATACACTA GGCTAAACAC CAGTGAACGT
39901 ATCAAAGAGC TTTTACAGA CTGTGGTCCC CTGGAATCCT CTTCTCTTTG
39951 TGACTACAGA AATGACATGA ATGCCAGTAA AATTGTCGAT GACATTCCTG
40001 ATCGTCCAGC AGGCATTGGA GTATATTCAG CTATATGCA GTTATGCCTG
40051 GCACTCATAT TTTAAATCAT AATGACAGTA TTCACTTTTG GCATCAAGGT
40101 AAGTGCTAAT GTGAGGTGAT ATTTGGGTAA TTTTGGCATG TTCAAAACTT
40151 ATATGTGGA TGAGAGAGGT TGTGTTTCA TAAATGACTG AAAAAAGTAC
40201 TTATCTTTTG AGTTTAATTT TAAGTAATGA AAAAGATAAT TCCTTAGCAT
40251 ATATTGTTGA CCATGTTATC TGTGCTATT TAACAAATTA CCCCCAAAA
40301 CTTAGCAGCT TAAGGTAAC ACTTATTTTG TTCTTGATAT TGAGTCAACG
40351 ACTTGGAAG GGCTCAACTG GGCAATTTTT GCTTGTGGTC TTTTCATATAG
40401 TTGTTATTAG ACATGGCGAG GGCTAATCAT CTCGAAAGCTT CTTTTTTTCTG
40451 TTTCTTTTTT AAAAACTGT TTTTGTGGAT ACACAGTAGC TATATATAGT
40501 TTTGGGGTAT ATGAAGTATT TTGATAGAGG CATGGAGTGC ATAATAATCT
40551 CAGGGTAAAT GGAGTATCCA TCACCTCAAG CATTTATCCC TTGTGTTACA
40601 AACAAATCCAA TTACACTCTT AATTATTTTT AAGTGTACAA TTAAATTATT
40651 GAATATAGTT CAAAGACTTC TTCATTCATG ACTAGCACCT AGGCTAAAAA
40701 AATTCAGACA CCTGGGCTCC TGGGATCAAT CACGCATACT GTGTCTCTTG
40751 TGCTCACTCC CGCTGTCTCT CTCTCTTTCT CTCGCTTCCT TTTTCTCTCT
40801 TCTCTGTGGT TTTCTAGGGT GGTGGCCTCA GGAATTGGA TTTCTTATAT
40851 TATAGCTCAG GATTTCCAAG AGGGCTGTTT TTAATGTAGC CAAAGAAGTC
40901 TTGCAGCGTG ACTTGTTTTA TTCTATTCAT TGAGGTAGTC ACAGAGGCCC

FIGURE 3, page 13 of 27

40951 GACCACATTC AGAGGAGGGA CATACACTTG CTGGGACAAG TGTAAGAGAA
41001 TTCATGATCA TGTTTTAAAA CCACTTTTAT TAGTTTCCTA TTGCTGCTGT
41051 AATAAATTAC CACAACCTAA TGGCTTAAAA GCCACACAAA TTTAATATCT
41101 TACAGTTCTG CAAATCAAAA GTCTGAAACG GATCTCACTG TGCTAAAATT
41151 AAGGTGTTTC TAGGGCATTC TGGAGGCTGT AGGAGAGAGT CTTGTTTTTT
41201 GCCTTTTCTG GCTATTAAAA GCTGCCAGCA TTCCTTGGCT CCTGGCTGTC
41251 TATTTGCATC TTCAAAGCCA GCAGTAGCTG GTCAAGTCTT TCTCTTGTCT
41301 CATCACCTCG ACCCAAACCTC TGCTAAATCT CCCTTCCACA TTTGAAAAAC
41351 CTTTGTGATT ACTTTAGGCC CACGCAGATA AATCAGAAAA TAATCTCCTT
41401 TTTCAAGGTC AGTTGCTTCG AAACCTTCTT TCTGCCACCT TGATTCTCTC
41451 TTGCCATGCA ACGTAATGTA ATCACAGGTT CTGGGAATTA AGTTATGGAC
41501 ATCTTTGATG AGCCATTATT CTGCCTCATA CCAGTATAGG GTATTAGCTT
41551 GAAAGGACAC TGCAGACTCA GTTAAATTAC TAGATCTATA AATACATGCC
41601 TTTTTCCATC AAGAAATTAA GGCAGCTGGG TCTTATGCCC TGGGACATTG
41651 CTTCTTTTGG ATTTATAAAA TAACAAAATT TGTTGATTAA TGGTCTATCA
41701 GTAAATATAA TTTCTTATGT GACTATCAGT GATATATATG GGGAAGCACA
41751 TATCAGCTTA TTTCTGTTCT TTAAATTACT ACCCCTGTAC TTCATGTAAT
41801 AGTATTTGCT AGTGATGATG TGCTTTTACA GATGTAAATT AATGTGGAAT
41851 AACAGCTTTG TTTCTACAAA ATTAGAGTGG TTTTAGTTTT TGAAATAAGG
41901 TCTCTTTTCT CTTGTCCTAA GTCTGTAGTC CACTGAGTAT CTAGAGTTAA
41951 ATAATAGAAA AGCCTGGCCA GGCGCAGTGG CTCACACCTG TAATCCCAGC
42001 TCTTTGGGAG GCCAGGCGGG GCAGATCACA ATGTCAGGAG ATCGAGACCA
42051 TCTGGCTTAA TGCCTGTAAA CCCCCTCTTT ACTAAAAATA CAAAAATTAG
42101 CCAGGCGTGG TGGCAGGTGT CTGTAATCCC GGCTACTCGA GAGGCTGAGG
42151 CAAGAGAATC ACTTAAACCC AGGAGGTGGA GGTGCAATG AGCCAAGATC
42201 ACACCCACTG CACTCCAGCC CAGGCAACAG GGCAAGACAC TGTCTCAAAA
42251 AATAATAATA AGAAGAAAAT AATAATAGTA ATAGAAAAGC CTAACATTTT
42301 TACCTTTTTT TCTTAGGGAA TCAAGTTAAA AGAGCTGTAA AAGCTCTTTT
42351 TCCTACAATA AGTAAGTGT GGGTAAATCC CACTTTCTC ACAGTCAGTT
42401 GAACTACAAG AAGCTGGAGG CAATTGGCAG GCCTTTGTTA AGTCCCACCT
42451 TTGACTCAGC TCTGGCTGAA GGATCATACC TGGCAAGAGA GTGTAAAACA
42501 CACTTTGATT TTTTCTATTG TTTATCCTTT TAATGATCCT AAGAGACTCA
42551 AGAGTACATG CCATCATTTT GTGTTTGGCT CATTTCATAT TCAGAGGAGT
42601 TTATTACTCT TTCAGTAGTT TGTTTGTTCG TTTGTTGTT TTTTGAGACA
42651 GGATCTCGCC TTTTGGCCCA GACTAGAGGG CAGTGTGCA GTCTTGGCTC
42701 ACTGTAACCT CCACCTCCCA GGTTCAGCG ATTCTCCTGC CTCAGCCTCC
42751 CAAGTAGCTG GGATTACAGG TGTGGGCCAT CACACCCGGC TAATTTTTGT
42801 GTTTTTAGTA GAGATGTGAT TTTGCCATGT TGGCCAGGCT GGTCTGGAAC
42851 TCCTGACCTC AGGTGATCCT TTGGGAGGCC TTGGCCTCCC AGAGTGCTAG
42901 GATTATAGGT GTGAGCCACT GAACCTGGCC TCTTTCAGTA GTCTTTAAAT
42951 GATCTTGCTT ATGGTGCTTC TTATCCCTGT TTATTATCCT TATTAAATTT
43001 AATCAATAAA TATTTTTCTC TTTTAAATG ATTCATATAA ATAGACTTAC
43051 CTGAGAGATA TAGGTTTCTG TCAGAGCACC ACAATAAAGT GAATATCATA
43101 ATAAAGCAAG TCACATAAAA GTCTTAGTTT CTTAGTGCAT ATAAAAGTTC
43151 TGTTTACACT ATGCTGTAGT CTTATGTGTA CAATAGCATT ATGTCTTTTA
43201 AAAAAGTAAT ACTTTAATTT AAAAATACTT GATTGCTAAA AAATGCTAAT
43251 AGTAATCTGA GTCTTCAGTG AATTGTAATC TGTTTTGCTT CTGTAGGGTC
43301 TTGCCTTGAT ATTGGTGGTT GCTAGAGGTA GGACTGGCTG TAGCAATTCT
43351 TAAAATAAGA TAACAGTGAA ATTTGCCGCA TTGATTGACA CTGCCTTTCA
43401 TGAAAGATTT CTCTGTAGCA TGTGATGCTG TTTGATACCA TTTTACCTAC
43451 AGTAGACCTT CTTTTCAAAA TTAGAGTCAT CCTCTCAAAC CCTGCTACTG
43501 CTTTATCAAC TAAGTTTAAG GAAAATTCAA AATCTTTTGT CCTTTTAAAC
43551 ATGTTACAAA CATCTTTACC AGGACTGGAT TCTACCTCAA GAAACCACTT
43601 TCTTTGCTCA TCCATAAGAA GTAACCTCTT ATACATTCAA GTTTTTTAAA
43651 TGAGATTCTA GCAATTCAGT CACATCTTTA GGCTACGCTT ATCATCTAG
43701 TTCTCTTGCT AATTCCACCA CTCTGTAGTT ACTTCTTCAA CTGAAGTCTT
43751 GAACCCCTCA GAGTCATTCA TGAGAGTTGG AATCAACTTC TTCCAAACTC
43801 CTGTTAATAT TGATATTTTG ACCTCCTCCC ATGAAACGTG AATGTTCTGG
43851 ATGGCATCTA GAATGGTGAC TACTTTTTGA ACATTTTCAA TTTATTTTGC
43901 CCGGATCAAT CAGAGAAGTT GTTATCAGTG GTGGGTTTCC AAGTTGTCAG
43951 GGGCGAACCA TACAGATCTT CAGCAACCTC AACTCTTGCC TTCTCAGAGG
44001 AAAGAATTCT ACGGAGGGAC ATAAGGCAGA AAAAGAGACT GAGGCAAGTT
44051 TTAGAGCAGG AGTGAAAGTT TATTATTAAA AAGCTTTAGA GTGGGAATGA

FIGURE 3, page 14 of 27

44101 AAAGAAATTA AAATACACTT GAAAGAGGGC CAAGTGGGCA TCTTGGGAAGA
44151 CAAGTGCCCC ATTTGACCTT GGACTTAGGG TTTTATATGT TGGCATACTT
44201 CTGGCATCTT GCATCCCTAT TCCATTGATT CTTCTTTTGG GGTGAGTTGC
44251 CCACATGCTC AGTGGCCTGC TAGCACTTGG GAGGGGAGTG TGCACAGTGT
44301 ATTTACTGGA GTTGTATGCA TGCTTACCTG AGGTGTTTGT TGCTTACCAG
44351 CCAAATGTCC CTAGGAGGTC ATATTCATAA ACTCCATGAT TTTGCCTCTA
44401 AATGTGCATG CTTGAGCCCA CTCACCCAAC TCCTGGGATC TTATCGGAAA
44451 GCTGCCGATC GCTAGTTTCA GGTGTTTCTA TCTATTGGAA GATGGCCTTT
44501 CCCTGATGCT GGCTGCAACC AATTATTACT TTAGAGAGAG AGCATGAGAG
44551 CTGTCTCACC ATCATCACCT GATGGTTGCC TGACATTCCT GGTGGGGTTG
44601 GGAGGATGCC TGTCTGCCCC TGCTCATGCC TGACTAGCTA CCTGCTGTAA
44651 CAAAAGTACT ATCTATGGTA GCTGTAGCCA TAGGAAATGC ATTTCTTCAG
44701 TAAAACCTAA AAGTCAAAAT TAGTCTTTAA AACAACATGA ATCTCCTTGT
44751 ACATCTCCAT CAGAGCTCTT GGAAGACCAG GTGCATTATT AGTGATGAGT
44801 AATGTTTTAA AAGGAATCTT TTTGTCTGAG CAGTAGGTCT CAACAGTGGG
44851 CTTAAAATAG TTAGTAAACC ATGCTGTAAA CAGATATGCT GTTATCCAGG
44901 CTTTGTTATT CCATTTATAG AGCACAGAGA GAGTAGATTG GCATAATTTA
44951 AGGATTACTT AAAAAAAAAAG TCTTTGATTA CTCTCAAAAA AAAGTCACGT
45001 CTCTCACTTT ATATCAACAG CTAAAAATGG CCAGGTATTG TGGCTCACGC
45051 CTGTAATCTC CATGCTTTGG GAGGCCAAGG CAGAAGGATC ACTTGAGGTC
45101 AGGAGTTAGA GACTAACCTG GGCAACATAG TAAGACCCAT CTCTACAAAA
45151 AAAAAAAAAA AAAAAAGAAA GCCAGGTGTG GTGGTGCACG CCTGTAGTCC
45201 CAGCTACTCA CGAGGCTGAG TCGGCCAGGAT CACGCCACGC CAAGAGACGT
45251 GACTTCTGCT TTCAGTTGTA CACTTAGAGA CCATTGTAGG GTTCTTAGTT
45301 GGACTAATTT CAATATCATT GGGTCTCAGG GAATAGGGAA GCCTGAGAAG
45351 AGGGAGAGAC AGGGGAACAG CCAGTTAGTG GAGCAGTCAG ACCACATACA
45401 AACTTATTTA AGTTCACCTT CTTCTATGGG CATGGTTCAT GGTGCAGTAA
45451 AACAACTGTA ACAGGAACAT CAAAGATCAT TAATCACAGA GCACTGTAAC
45501 ATATAATAAT AGTGAAAAAT TTCAAAGTAT TGAGAGAATT AGCAAAATAT
45551 GATACAGAGA CACAAAGTGA CCACATGCTG TTGGAAAAGT AGTGCTGATG
45601 GACTAGCTTG ATGCAAGGAT GTCATAAACC TCAATTTGTG AAAACTGCAA
45651 CATGTGTGAA GCACAGTAAC ACAAAGCATA GTAAACAAG ATATGTCTGT
45701 ATATCAGTCA AAATATTGGG CAACTCTGAT AAGTTTGTCC ACTTAACATT
45751 GTACCACTTA AGATGAATAG CATCTACCAT TTCCGTCATT TGTAATATA
45801 TAGGAGGACA TAATCACATA ATCTTGAAGT AAAAGACAGT GCTTAAACT
45851 GAATCAGTTA AGTTTTATGA AAAATACTTC ATATTGTACT TTTAAAAATA
45901 TATATTTTTT AATTTCAATA GCTTTTGGGT TACAAGTGGT TTTGGTTACG
45951 TGGATGAATT CTATAATGGT GAAGTCTAAG ATTTTACTGC AACTGTCACC
46001 CAAGTAGTAT ATATTGTATC CAGCATATTG TCCTTTTTTT TTTCTTTTTT
46051 TTTTTTCATT TCACCATGGA CTAATGAAAA TTTTGTTAGG GACTGACATT
46101 AGGGCACCTT TGAGCTACCT TGAGCTAAAG GAAATAACCC TTGAATTTTT
46151 TTTCTGTTGG CCTAGAGAAT GTGGTTTGTG TTGTAAGTGA ATTCATGGGA
46201 TTGTTAAGGT ACAAGATTTT GCTTTAGTTT TATTTGTACT AGGATTTTGC
46251 TATATTAATA CAATGTGAAA AGAATCAAAA GTGTTAGAAA TAAATGCATA
46301 GAATGTAAGT TTCAGGCATG TGAGTAGAGG ATCTCTGCTC CATAAAGAGT
46351 TCTGTTGTTG TTATAGGTTT CATCAGGCTT GTTCATCCCC AGCATGGCCA
46401 TTGGAGCGAT CGCAGGAAGG ATTGTGGGGA TTGCGGTGGA GCAGCTTGCC
46451 TACTATCACC ACGACTGGTT TATCTTTAAG GAGTGGTGTG AGGTCGGGGC
46501 TGATTGCATT ACACCTGGCC TTTATGCCAT GGTGTTGCT GCTGCATGCT
46551 TAGGTAATAT GGCTGTGTCT GCCTGTGTGT GGATGTTTGC AAGTCTGAGA
46601 GAGCCAAGAG AAAGTGGGAC ACATTCTTGC TTAATTGGTG GGCGGATTGG
46651 TTGAGTAAAG GAGGGTGCCA GGAGGAGATG TTTTAACAGA TAAGAAACAG
46701 TAGTACTATT AGGGTATTAT ACAGTACCGG TTTTCTGTCT TACAACATTT
46751 GTTAATACAA GAATTTAATG GCATTAGCAT ATTGTAATAT AACTTAATAC
46801 ACTATGGCAG AAGCCATCTA AGTACAACAT AAGCTTAATT TGAATCCTGA
46851 CCAAAGATGT CTTTGATTCT TTCATCGTTA AGGATCTTGG CTTACCTATA
46901 ACAACTATAG CATAATACCT AAGATTAGCA TTGCAACAGA GTTTCAGAGT
46951 AGGTTTACTT TGTTCTGAA ATGATTTATT GTTAGCCTTA GTAAAAGATG
47001 TATTTACCCA TGCTCCATCA TCTAAGGTAT ATTTGTAACA AAATGAGAAA
47051 AGGTAACCTC ATTTTAATGA GAAGAAAAGC AAAATACCTA CATTAAGTAC
47101 TTGAGTCTAT TTAATGTCTG TTAGGGCAGG AAAAAATGGT TATTGCTTTT
47151 CATATTTAAA ATATCAGCTA CACTCTGGTG ATAATATTAA TGGTTGCCAT
47201 TTTGACCAGT TTTGTTTAGT GAATAAAAAAT TATGTGATTA TTGATCTTTA

FIGURE 3, page 15 of 27

47251 AAAATGTAAT ATCAATTAAA AGGAAAGGAC AGACTCATTT TCACCAAAGT
47301 AGCAAGTATT TATTAAATGT CCACTTTCTT TTTAGCATTG TGCTAGATAC
47351 AGTGCATAAT ACAAAGAA CATGGACCCA ATCTCGACTC TAATCAAGTT
47401 GAGGAGACAA GATGAACACT GAGAATACAA TAGTGAGGAA TACTAACAAA
47451 TATATACAAG GTTAAAAGAG TCTAAGTATG GTAGGAATAT AGGGGAAGAA
47501 AGAGCTGAAG TACTTCAGGA AGAGTAGAAC ATGAGGCTTT ATTTAAAGA
47551 TTAGCAGAAT TTAAGGAAAA GGTGACTTTG TTGAAGATTA TAATGTGAAG
47601 ACAAAGGAAC GAGGATGGGA ATAAATTTTG TATTCATGAG GCTTTGAAGA
47651 AATTGACTCT AGAGAGTATA TTTTGGGTAC TTTTGGGAAA TGAAGTTGGA
47701 TTAGTGAGAA GGAACAGATT ATGAAAAGAC AAGAAACCTG ATTAATGTCA
47751 GGATGATTTT ATATTTGAAG TTGGTCAGAT TTATGGCAGT CCTGGCTTTG
47801 CCATTTTTAG TTTGATGACT TTGAGAAAGT TCCTTCTTGA AGTTTTAATT
47851 TTCTGTATAT AAAAAGTAAT AACACCTGGT GATCTGCTAG GTTTGTTTTG
47901 AGGATTATAT GAGATAAAAT GCATGCAAAA CTGTTATAAT AGTGCCTGGT
47951 AAAATAAGTG CCTAGTTTTA AAAACAAGTC TTTGTAAACT GCTTAGGACA
48001 TGCCTGGTAT AGGGTAGGTA TGTAATACAT AGTAGGTAGG ATCTGTCTCC
48051 TTGCTATTTT TAGGTAAAAA AACAAAAGGA AGAGCTTCAG CTTAATACAG
48101 TATGAACCTGA CGAGCCCTGG TAGGTTTTTG AGCAAAAGAG CAACACAGTA
48151 AAAGTAGTAC TTAGGAAAGA TTAACAAGGG AACATGGCTT ATACAGTGGT
48201 AATGGGGCCT GGAGTCAAGG AGGTAAGATA AAATGGTATT ATAATTAGG
48251 AATAGCCAGG CACGATGGCA CATGCATGTA ATGCCAGCTA CTGGAGAGGC
48301 TGAGGTGGGA GGATCATGGG AGTCCAGGAG TTTGAGACCA GCCTGGGCAA
48351 CTGAGTGAGA CCCCAGATCC TAAAAAATAC AAAGTAAAAA AGGAATAAAG
48401 TCATGAGGGC TTGGACTGGA TTGATAACAG TGAGAATACC GAGAAAGGGA
48451 CCATAGGCAG TGTGAACGCA GCTCACTGCA GCCTCAAACC CCAGCCCCAA
48501 CGAGCCTCCC ACCTCAGCCT CCCAAGTAGC TGGGACCACA GACATACACC
48551 ACCATGCATG ACTACTTTTT TTAGTTTTTA CTTTTGTAGA GACAGGGTCT
48601 CACTGTATTG CCCAGGCTGG TCTCAAACCT CTTGACTTAA GTGATCTTCC
48651 TGCCTTGGCC TCCCAAAGTG ATTACAGGCA TGAGCCACAG TGCCTGGCCC
48701 AAATAGTTTT CTGTGAGTGA ATATTACTTG CATCGTTAAT GTAAATCAA
48751 GGCATCAAAG TATTTTACTC TTTTGAAGA AAATTTAGAG GAGAAATTTA
48801 TTATATTAAT ATTCTACCCA TATATGAGTT TAATTTGTAA ATTGTAGCAA
48851 AGCATGATGT GCTTTACTAA ATTCCCTTAT AATTAGAATA AGCTTTTATA
48901 AGGGTGAAAT TATGTCTTTG CTACAGCACT AAACCAAAAT GGCAAAATTG
48951 TTTTAGTCGG TAAGCTTTGC TTTTAAAAA TATGAAATAA ACAGGTTTTT
49001 AAAATGTTAT TTTAATAGTC TTCTCTGTTA TAAACAAAGA AAATTGGTGT
49051 TTCTCTAGAG CTTATTAAAA GTAGTGATTA TTGTCTTAAA AGAGGAGTAG
49101 CAGTTTTAGA TGCTAATGCT TTTCCCTGAC TGAGTTCTAT TTGCCATTTA
49151 GTTTTAACTG CCTAGTGCAA AAATTCTAAT AAAATGTAAT GATGAGGATC
49201 CTGTCTTCC TGACCAGTGG GTGCTTACTT TTTTCAGGTG GTGTGACAAG
49251 AATGACTGTC TCCCTGGTGG TTATTGTTTT TGAGCTTACT GGAGGCTTGG
49301 AATATATTGT TCCCCTTATG GCTGCAGTCA TGACCAGTAA ATGGGTTGGA
49351 GATGCCTTTG GCAGGGAAGG CATTATGAA GCACACATCC GATTAAATGG
49401 ATACCCTTTC TTGGATGCAA AAGAAGAATT CACTCATAAC ACCCTGGCTG
49451 CTGACGTTAT GAGACCTCTA AGGAATGATC CTCCCTTAGC TGTCCTGACA
49501 CAGGACAATA TGACAGTGGA TGATATAGAA AACATGATTA ATGAAACCAG
49551 CTACAATGGA TTTCTGTCA TAATGTCAAA AGAATCTCAG AGATTAGTGG
49601 GATTTGCCCT CAGAAGAGAC CTGACAATTG CAATAGGTAC CCTTCAAAA
49651 ATATATATAT GTATATATGA GATGGATTTC TGGAAGAAAG GAAAGCAATA
49701 AGCAGTAACA TTTAATGGGT CGGATTTGTG GGGGCAAGGG ACATTATTTT
49751 ATGTCCCTTA ACATCTTCTG TTCTTTAAGA AAGGAAGGTA TGCTTCAGTG
49801 GATGATTTTC TGCTATATAT CACAAAATCT GTATTTCAGG TTGTCTTTT
49851 GATCCGGCAT GTACCAGAAA TTGGAGTCAG ATTATTTTCC CACTCAGATA
49901 AGCCTAGATA AGTTGATCTT GGTATTTCOA AACAGCATGT AATATAAGAC
49951 CTTAGCTAAA TGCATTGAGT CAAATACATT CTTGTATTTA ATAAAGTTGG
50001 CTTATTGGAA TACAAGTTAT TGAAAACTC ATCTTCATCA GTCTCTTTCA
50051 TATTAGAATA AACTGTTTTT GCTTTATCAG TCTTTGGGGT TAGAATTATA
50101 ATATTAATTT ATAATATCTG ATTTAAAGTG ACAATCACTG AGATTTTTAT
50151 TTCTGATCAA ATGCCAGGTT GAAAAAGTAT AACGTATCAG TCCTGTTGTG
50201 TTTTATGCAG ACTTTCCTGA AAATACTGTT TAAAGGTATT AGCCATAGTG
50251 TATTTCTTGG AGATAAATTA AACTTCTAT AGTTCTGTTT CTCTAAAATT
50301 TGTCTTCTC TTTACCTTAT AGTCCCGCAG TATTGATGAG GAGACCATTA
50351 AGACTTAATA TTTTTTTGAC ACAATCTTAT ATCTCTTCTT CCAACCCCTA

FIGURE 3, page 16 of 27

50401 AAAAGTGACT GAGGATAGGT ACATCAAGCC ATTGCTTTGT TACTCCCCAG
50451 GTTTTAGTGC CAGACCCTGA ATGGAAGTGT CAAGCCTTTG GCCTGTCTGA
50501 AAGGTCATTCT CTGTGAGCAT ATCATCTCCC TTCCAGCTTA CCTCTGTGGC
50551 CATTGCAAAA GGATTTAAAA ATAATTTTGT TGCCATTTGA ATGGCACAAG
50601 ACCAGACAGT GTATGTGGGG GAGTGTCTCT CAAATCAAAC TGGAACTCT
50651 TTAATTTGTA AGAACCATTA AGCAGAGAGA GAAAAAGAA AGGAAAAGAA
50701 AAAAGATCCT ACAGAGAACA CCCTGTTTCA TTTGGGAACA GGCTACAGCT
50751 TTGGATTTT CAAGGCCTAG CATTCCTATC ATTCTAAAT TTACTTAGCT
50801 AATACAATAG TAGTTGCCAG AGCTGATGAC ATAGTATTTT GTCATGCTTG
50851 GCTCCGTTCA AGCATTTTAG TTTTCTAGCC ATTACCATGG CTAGACCCAG
50901 TCAAAAGAAT TTTCATTGTT TAAGATTCCC ATTATCCTAG TTTTACTAG
50951 TAGCCAGCCA AAGAAAAGAA AAAGGAGGTC AGAATTTCTG TATTTACATA
51001 GAAATTTAAG GGGAAAAGGC CAGGCATGTT TTTAAAGTGT GGAAATTAAG
51051 AACTATTTCAT TATCCCCTG ATTGTGTGGA TGTGTTTTT AAAGTTTTGT
51101 TACTGTCTTG AGAGAGAGAA TATTGAGATA GGACATAATG TTGGTTTAAG
51151 GGAATGAGGG TACTTTCTGT AGGTGAGGTG CCAAGCCATG TCATCAGAAA
51201 TGTTAGTCAC ATGACTTCT AAGCACACCT TAAATGTTTT ACCGTGTATG
51251 TTTTGTAAA GTTTTAAAT TTTAACTGGG AAAACAGAC CTGTATATTA
51301 AGTTTTATAT ATATATATAA ATTTAAAT ACATATATAT GTTTATATAT
51351 GTAACTTTTA TATGGGAGAG ATATATATTT CTATATCCTC TATAAAAAA
51401 CATATCTATA TATGAAAT ATGTACGTAA ATGTTAATTT ATAATTAAT
51451 ATATAAATAT TAAACATAAT ACATTATATA TATAGAAAAC CTAGTGACA
51501 GATCTGTATA TAAATTAATA ATGTATGTGT TATATATAGT TACATCATAT
51551 AATACATATA ATTGATATAT ATAATGATAA ATACTTTATT GAAGGATGAA
51601 AAAATTTCCA TGCTGTCTCA TAAAATAAGA TGGTTGACAT ATGCTAACT
51651 AGATAGATTC TCCTGTTTCA TACTAAAGCA GAATGTTGTA AAATATTAAA
51701 TCCAAATGAG ATGTCTCAGA TTAAGGCCAT TTCACAGGA ATGCTGAGAC
51751 TTTAAAAAAA AAAAAGTCT GAGGCTGGGC GTGGTGGCTC ATGCCTGTAA
51801 TCCCAGCACT TTGGGAAGCT GAAGCAGGTG GATCACTTGA GCCCAGGAGT
51851 TTGAGACCAG CTTGGCCAAT GTGGTGAAAT CCCGCTCTA CTAAAATACA
51901 AAAAAAATAC ATGGGTGTGG TGACGCATGC CTATAATTCC AGCTACTTGG
51951 GAGGCTGAGG CAGGAGAATC ACTTGAACCT GGGAGGTGGA GATTGCAGTA
52001 AGCCCCACCA CTGCACTCCA GCCTGGGCGA AGAGCAAAAC CCTGTCTCAA
52051 AAAAAAATAA AGCCTGAATT ATATCAGCAA ATGAAAAC TGAAATGTTGT
52101 CTCTGTTTCA GAGGCCCTTG AATGAATAGC ACTAAAAATA TTTTAAAAA
52151 AATGAAGAAA ATGAAAATTG TAATGTTTCT TATTTAAAG GCCCTGAAT
52201 GAGTAGCATC AAAAAATATT TTAATGGGA GGCCAGGGTG GGAGGTTTGT
52251 TTGGCACCAG GAGATCAAGA CCAGCTTGGG TAACATAGCA AGACCTTTGT
52301 CTCTACCAAA AAAAAAAT TGGGTGTGGT GGTGCCACCT GTATTCTAG
52351 CTACTGGGAA CACTGATGCA GGAGGATCCC TGGGACTCTA GAGTCCAGAG
52401 TGAGACCCTG TCTCTAAAC AAACAAACA AAAAAACTG TATTTATGTA
52451 AAAGTAATAC TTGTTTTTAA ATTTTATTT ATTTTAAAT GATAAAAAAT
52501 GTATGTATGT TTATGTATG TATATATGT GGAATGGTTA AATCAGGCTA
52551 ATTAACCTAG ATTTTGTGT TGTGTGGGGA GAATATCTAA AATCCCTCTC
52601 CTTAGCAGTT TCCAAATGAA ATGAAAGAAT AAAAGTGATT TATTTTTTTG
52651 AGACAGCATC TCACCCTGTT TCTCAGGCTG GAATGCAGTG GCACGATCTT
52701 GGCTTACTTG ATCCTCGACT TCCCTGGCAT CCGGTGATCC TCCCCTTCA
52751 CTCTCCTAAT TAGCTAGGAC TACAGGCATG CGCCACCATG ACTGGCTAAT
52801 TTTTGTATTT CTGTATAGG CAAGGTTTTG CCATGTTGCC CAGGCTGGTT
52851 TCAAGCTCCT GGGCTCAAAC GATCCACCTG CCTCAGCCTC CTGAAGTGCT
52901 GGGATTACAA GTGTGAGCCA CCACACCTGG CGAAAAGTGT TATTTTTTTA
52951 AATGACAAAT TTAAGTCAA GAGATTGAAT GTTCACTTCT GGTACTTTGT
53001 ATATAAGAGA AACATTCCAT TAAATAATTT TTTAAACATT TCTAAAATTA
53051 CATATTTTGT CATTAATGT TTAACAATC AGTATAATTT CATTGATACA
53101 GTGTTTGTTA TTTTGTCTG GTTTAAGATT GATAATTGGG GTTAGTTTTA
53151 ATTCAGAAATG TTATCTATT TAATGTCACA CTTTATGCT TTTTATTTTG
53201 TATATCTATT AATGAATTAT TTTAGCTATA GTTATTACTG TTTTAGAGAT
53251 GAGGTCTTCT ATGTTGCCA GGGTAGACTT GAACTCCTGG GCTTCAGCAA
53301 TCCCCTCTC AACCTCCGGA GCACATGAGA TTAGAGACGT GTGCCACTGT
53351 ATCTGGCCTG CTGTAGTTAT TTTAATTTCT TTTGCTTTT AACTTTTATA
53401 CTAGAGTTAG AAATGATTTA CAAACCCTAT TGCAGTTTTA GAGCGTTATG
53451 AATTTGACTA TATATTTCTT ATAACAACCT AACTTCAGTT GCTTACAAAA
53501 ACTACAGAGT TTTACTCCCC CGTCCACATT TTATACTATT GATGTCACAC

FIGURE 3, page 17 of 27

53551 TTTACATCTT TTTATTTTGT GAATCCATTA ATGATACTTC TGGTAGTTTT
53601 TACACTCCAC TATTCAGTTG TCAGACACCA TTCAGTTGTT AGATTGTTAT
53651 GAGCTAAAAG CAACCTTAATG GGTATTTTTT AAAAATCATT TATGTCAATT
53701 GCTAATGGAC TTCTTTTCTA TGCCATGATC ATGCTTTTTT TATTTTTGAG
53751 ACGGAGTTTC ACTCTGTGTG CCTGGGCTGG AGTGCAATGG CGCGGCCTCA
53801 GCTCACTGCA ACCTCCGCCT CCTGGGTTCA AGCGATTCTC CTGCCTCAGC
53851 TGGGATTACA GGCATGTGCC ACCGTGCCGG CTAATTTTGT ATTTTTAGTA
53901 GAGACAGGGT TTCACCATGT TGGCCAGGCT GGTCTCGAAC TCCTGACCTC
53951 AGTTGATCTG CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGACGTGA
54001 GCCACTGCGC CTGGCCTGAT CATGCTTTTA AGGTGGTTGA GTAAGTACTA
54051 GTTGCTGGGG CTTTACTTAG TGCCCTCCTA CTCAAATGTG TTAGAACATA
54101 GTTAAGAAGG CTGTAGTGTT CAAAAGGAGT AAAAAGCAGT GCAGTGTTTG
54151 CAGTAATATC TGCTTCTCAA TTTAGGACTG ATGCTTATTA TGGCTTAAAT
54201 GTTTTTGTAG TAAAATTTGT ATTCAAAAAA TATATTTTTT TTTCTTTTTG
54251 CGACAGAGTC TTGCTTTGTC ACCCAGGCTG GAGTGTGGTG GTATGATCAT
54301 GGCTGACTGC AGCCCTGACC TTCCGGGCTC AAGTGATCTT TCCACCTCAG
54351 CCTCCCAATT ACTTGGGACC ACCAGCATGC TTGGCCGATT TTTTTTTTTT
54401 TTTTTTTTTT GTAGAAGCAA GGTTCCTTA TGTGCCAAG GCTGGTCTTG
54451 AACTTTAGGG CTCATGTGAT ACTCCTGCCT CGGCCTCCCA AAGTGTTAGG
54501 ATTACAAGCC TGAGCCACCA TGGCCGGCCA AAATATTTTC ACTATAACAA
54551 ATATCATATC TGTATATACT CAGTTTAAAT ACTAACTCAA AGTAGAAACA
54601 TAAAGCTGAA TGACTATTTT ATTTTCAGAT TCTCTCCATT GAGTTTCCTT
54651 CTCCGTCTTG TGTGATCTCT GAACCTTTCT CCATCTTTGC CACTTCTTGT
54701 CTAGCATTTT TTTTTTATCA GCAGTTTCAT TCAGATTTTT TTTTAGTTC
54751 TTTCAACGGT GGAGTGGAAG TAGGCAGCAG GACAGAAGAA CTTGAAGCAG
54801 AGCACACTGG AGAGGAGAAA TTAACAAAGC CTTTATGAAT AAAACAACCC
54851 CCAATATCA GTCTGTGTGC ATTATGAGCA TAATTGTACT TTCATCTCAT
54901 CTGTAATGTT CATGACTTTT CTAGAAAATT ATACTTTAAC ATGAGAAAAG
54951 AAAAGAACC AGCTAATTCA TAGGGATGGA GGACACAGCA TAGTCAAAGC
55001 AAGAATGAAA AGTCTTTTAG TGCCACCTCC AGTGCAGAAT AAGTAACATT
55051 CAGCAGAGGC AGGTTTCATT TGATAATGGA TTCCTATAAT AAAGTGCCT
55101 CAGAATTTGT GCAGGTTTTA AAATCCCGTA TTCCAAACCC ACTTCCTTAG
55151 CCCCAGTT AGAAAACAGC TTCAGTAAAG AAAATGTGAC GATGATATAA
55201 CTTTACCAA AAATAATTTT TTTCCATGAA GATGATATAT TATTGTTGAC
55251 TTCTAATTCA ATCAAATATA AACAATTGCT AAATGGCTTT TCAGTTGACT
55301 CCTTCTTGG TTAAGGAGAA GATAGGAAA AATGAAGGGA TCAGAAAGTCA
55351 TAGGATACAT TAATTTTTTT TATCTCTGAA TAAACAGGTT GCCTACTTAA
55401 AAATCTATCA GTTTAAAAGT GTTGGTCTCT TCTCTCTCTT TTCAGAAAGT
55451 GCCAGGAAAA AACAAGAAGG TATCGTTGGC AGTTCCTGGG TGTGTTTTGC
55501 ACAGCACACC CCATCTCTTC CAGCAGAAAG TCCTCGGCCA TTGAAGCTTC
55551 GAAGCATTCT TGACATGAGC CCTTTTACAG TGACAGACCA CACCCCAATG
55601 GAGATCGTGG TGGATATTTT CCGAAAGCTG GGACTGAGGC AGTGCCTTGT
55651 AACTCACAAT GGGTAAGTCT GGTACCACAG GAATCAGTTC ACTTGCTAGA
55701 ATATAGGATC CTTTTTAGTG GAATCTATAT AGTTATTAGG GGAGCATGTG
55751 AGTCAGCTCC CAGGTGGGAA AGTCTGTCTT ATGGTATAGT CACAAATATA
55801 GGATCAGTCA ATCAAATTTT ACATTTACTA AGGAATAAGA AAGATGTCAT
55851 CTGCCTGCTC TTTGCCAAAC AGTGACATTT GTAAATAATA CCTCAAAGTT
55901 GGAAAGAGG TGCTGAAAGA TCTCCAGCAT GAAAGCATGT TGAGCTTAGA
55951 GTGCTTCTTT TCCTAGGGAA GAGTGGACCT AACCTGCATG GAGCACTGCA
56001 AAAACCTGTT TTATTTTTGT AAATGTTTCA TTTTLAGTAT ATAAATTTCT
56051 AGTACAATAA TAAGTTTCTA GATATTTTGC TATTTACTCT TTCAGCCAAT
56101 ATTTGATTTA TCATGTAATG AAGGAAAGAA TATATACTTA AATGAAATTT
56151 GTAAATGAGC TAAAAATCTC CTTTAAACAA TGCTTTGTTT CCTTTTGTCT
56201 ACCTTTCTCT ATACACAAAT CTTTATATAT TATATAACTG CTAAGGACAA
56251 ATAAATACTC ATGTATTTAA AATGTATACA TTGATAATTT ATTTTTCAC
56301 CTTTTTACACA TGAACGCCA GTGTTTCTCC ATTGACAGGA ATATAGGAAA
56351 GAAACAGATG TCACGGGGT TGTGGAGACC TTAATGCACA GAATTGATTT
56401 AGCAAATACA CTACTTCGTC ACCACTGCTC TCTTTTCCTG GACCTGGGAT
56451 CTGTTTCTCC ACATTCCTTT CTTAGGACC CTTCAFTTCC ACTATATATT
56501 CTTTCTGTGTT GAACTTAAGA ATGTTGTTTT ATCCGAAGGC AAATACCAAA
56551 AAACAGAGGG TATTTCTGGA TTATGCATAA ACTGGATGGC TAATCCTGAA
56601 CAGCGTAAAG CTGGTTGAAA TTCTAAACAG AGAATCATAG CAGTTTTTTG
56651 TTGTTTTTTT TTTTAAACAT GTTGTAGAAA ACACATTGGT GACAGAATAC

FIGURE 3, page 18 of 27

56701 ATGACTCCTG TCCAGAGAAA GGAGAGAAAA AGAACAGAAA GGAAGGAAAT
56751 TTGTTTTATT AACACCTTCA TATTTTCTCA TTAACTTTG CAGGACCTCT
56801 GCAAAGTAGG TAGTTATATC CCTACTTTAC AGATGTAGTA ATTAAGCTC
56851 AGGAAGCTTT AATAATTTGC CCAAAGTCAT GTGGTGAACA AGTCATGGTT
56901 CAAGGAATCA GACTGTCTTT CCTACTTTAA AACCCAGCCT CTTGCTACTA
56951 TTTTGCCTG TAAGTGAAGT ATAGAAATCC TCTTTCTTTG TGATTTCTTA
57001 AACTACTAAA ACATTTTCTT GGCCAATATA TTAGATTGAG TTAAGAATAG
57051 AAATATGAAA CTAGAGAATT AGATCTATGT TTAGTGTTTT TCACTGCGCT
57101 AATTAATAA ACTCTTTAGG AATATGAAGT AAATCATTAA AGAGATAAAG
57151 CCCTTAAAGG CAGGGAGTTT AGAATTATTA AATTCTAATA ATTTAGATAC
57201 TGATTGGAGA AGAGATGTAT TCATAAGTTA TTATTGTTAC TATTTGTCTT
57251 TGTGTAATAT TGTTTGATTA AATGATGGCA CCGACTTCAT TAAGTTTAAA
57301 AACTCAGTAC TAGTTAAATG GGGCAACTTT TCATAAAGCT TTGCTAGTCC
57351 TTAGGCCCTT TTATTTGTTA AATGGCTCAA CTGGAACCTA AGCTGAGTTG
57401 TTACAACTA TTATTTGCTT CAAGTTGTTT TCTGTTCTTG GCATGGCTTT
57451 TTCTTTTGTG TACTGACAAA TATAAATGTT ATTCTGTTGA GTTATGGTTA
57501 AAGATGAACA CAGAAGCTGT AGGGATTAAT TTTCATATTT CAGTTTGTG
57551 ATTAATTCCC AGGTATTTGG CAGCATAGAT ATTAGAAAGG AAAATATTTA
57601 AAAGAAAGTG TAAAAATAAC GAAGTGATA GAGCGAGGGG TGGATAGCTA
57651 ATTAATAATT TGTCTGGTCC TGCCTGTTCA TATGAAAAAA GGGGTTGGAC
57701 TTTCTTCTAA GGAATATAT TAAATTGCTT TCATCATATT TTCCTTATTT
57751 CTGCTGTCA AGGAAAAATA ATTGATACAT ATATGGGGAG AAAAGAGATC
57801 ATTTAGGGAA GTGGCTCATG GGACTTTTGG TTTTGTGTTA AGTGTATTAG
57851 GAAGTCGGGT GTTTTTTTTC TCACTTAAAT TATTTAAAC CCAGAAAAGA
57901 AATGATATCT TCTGGTTTTT AAAGGAGACC ATGAAGTTCT GCATAGCTAT
57951 CATTGATGTG TAGTTCATAC TGCATTTTTA GAAGTGGAAA ATAGTTATTT
58001 GGAGGAAGAT AACAAATCTG GAACCTTAGG TGCAAGGAGA AAAAGAATAG
58051 ATGAAAGGGA AAGATGTTTG TAAATTATAA AAATTTCAAT TAGCTATTGG
58101 TTTTCTGCAC TTTATATTTT AACTGCAGAA TTTTCAAAA TCAGTTAATC
58151 TTGGTGGAAT TAGCAGGATG TTAATAGGAG TGACTCAGAA AAAACATTT
58201 TGTGACTGTC TAAGTTTGGA AAGTATTGGA TTAATAACAA TTGAGGTTTC
58251 TTTACTATGG AACTCCTCAG AACTTATAAT ATGTTGATAT TCTTTGATTC
58301 CCAGATGAGG GGATGGGTAA TAGGATACAT GGTTTCCAG ACTTGTGTTGA
58351 AAATGCAACT ATTTTGGGTG TGCAGGGAAG GATATAGTAG AACTCATGGG
58401 AACTGGTGTT TCTTGGAACA TGCTTTGGAA ATGCTGGGTT ATGCCCTGTT
58451 AACTCTTACA TCATTAGTTT TTAGCCCAAA AGGAAACAGC AAATAATGTT
58501 TTATATGAGC CACATTTTGC GTTGATTTTC CTCCACTCT GTAAAATTAC
58551 TAAAGCAGCA CTCTGACTTT ATTATGCTCA AATCGCTCTT CTCCATTAAT
58601 GTGTGTTTCT CCATCTTTTA GGGTTTTTAC TTTATAAATA CAGAGATTAC
58651 TGTGTAATAA TCTAAATTG CCACTGGGTC GTTATACATT TGTAACCTTC
58701 CTCACAGTAT ATTTTGTGAT TTGGCAGAGT TTACCAATAT AGATGATACT
58751 AACTGAAATT AATCATCTG TATAATTGGA TAGAAAAGCA TGAGTAAGAA
58801 TTCAATTGGT ATTATATTTA ATTAATTGCC AAGATTTTCA CATTTCCTGA
58851 CTACAACAAT AAAATCAAAT GAATTGATGG CTTAAAAAAA AGAAATCTCA
58901 AATGTTTAGT CAATGAAGAA CATCTATTGA ATGAGTGAAT GTTCATTATA
58951 TATAGTGCAT TTTCTGAGCT TTTTGGAGG GGGAGTTGC TCCCATGCTC
59001 TGAGAACTTT TAAGGATCGA TACATTATTT TTAACATAAT AATGAGAAAA
59051 CATGAGCAGA GAACCCATTT CTGTCATTCC CATTCTCTAT CCTCCTGCTC
59101 CCCACCTCC CACCCAGCC ATCAAGCTAA GTAACATTTT TACACCTGGA
59151 CGTAGCTATA GGAACAGGCT ACTTTGAAGT CTCCTAGTGA CATCCTTCAA
59201 GTCTGAATGT TCAAAGGCAG TTTAACAGGG AGGTTGACTT AATGAGATCA
59251 TCAAGGAAAT GTCCAGTCAT CCTGAAGGGT ATTTTGGATG GGCTTCCAGA
59301 ATTTAAAGAT TAAAGTTTTT TTAAGGTTT TTTATTTTCA CTGTTTATAT
59351 TGCCACATTA ATTTCCATTA TAAAACCAAG AACCATAGTT TTGTTTAAAT
59401 TAGCAATCTA ATTATTTTCA TGTATCCTCA TTATGAGAAT TTATGTCCAT
59451 CACTTTGCTT GATGTGATAA CAGTGACATG CTAAATGAGA AACAATTGTT
59501 ATTTAGAAAA AAATGCACAA AGTGAAAGTC CTTTTAATCC CTAATCATAA
59551 ATACATTTTA TTAGCTTACT TTAAGAAGTG GCAGTCACAG CTCCTGAACA
59601 TTAGGGAGTG TTTCTTTTGG TCAGCATAT TATTATTAGTG CACATTGCCT
59651 TTAATTTTAA TTTGAAATTA TAGTAAATC CACGGGAGTT TTTAAGTCTC
59701 CTCACAGCTT TTTGCTACCT TTTACCAAG GTAGATCCAG ATGATAACTG
59751 CTGTGTTGCT ACATCATAGA AATTAGAAAA ATATTTTCCT CTGAGGAAAG
59801 AACATTGTAA ATGAAACTCT ACATATCAGA GGTCTATAGC TATGTATCAA

FIGURE 3, page 19 of 27

59851 TATTAAGTTT CTTTTGTACT TTGCTTTGTA GTCATCTTCA TTCCAAACTT
59901 TCATAATTAT TATTTTACT TAAAAAGAA AAATAACCCA CCAATATTGA
59951 AGATTAGTAT TGTGTCACTT TTGAAAGTCA GTAGAATTTA TGCAAAAGGA
60001 ACCTGGAAC TTAATCATT TTGTTTTAT TTTCTAAAGT TCATGAGACT
60051 CATTCTTATG GTTCATGTTT TTATTTTTTC TCTCATTCTT TATCATTATG
60101 ATTGGAAACT CTTTTAATTT AATTCTCAC ACAGTTATTA GCATAATAAT
60151 CTGTTTCAGG ATTGTCTTGG GGATCATCAC AAAGAAGAAC ATATTAGAGC
60201 ATCTCGAGCA ACTAAAGCAG CACGTCGAAC CCTTGGTGAT TAGATATATC
60251 AGATCTCCTC ATTAGACACC TTAGAAGTCA GGAAGCATGA AACTTGTGAA
60301 CTGTTGAGTT CTGTCTTCC CAGATATCTG CTGAACAAAA ATATCCTACT
60351 ATGCTGCCAA TTACATTGT ATCTGATAAA ATGTGTCTGT AAGATAAATT
60401 TAGATATGTG TAAAATCCCA TTTATAGAAA GTAAGCAAAA GTTAACATCT
60451 CTCATCAAAT CATTCAATAC AATTTCAGAA CTGTAAACAG TTTGGTAGTG
60501 GAATAAGTGA ATATTATTGG ACATCTTAA AGTGAATATG GCAAATCTGT
60551 CTACCTCAGT GGATACACCG GTCTCAGAAG ACACCTGACT GGTTAAAAAT
60601 GTCTGACCCA TCCCCGCAAG CCCTTTTTTT TTTTTTAA TGTTCCCGA
60651 TCTTGTGGTA GTCTTATGGT AAATCTAAGC TCCTAAAGGA TTTTAAAGGA
60701 CTTAGCAAT TAGAACTGCT TACAGTTAAA TGGATTTTTT AATGGGCACA
60751 CTAAGTAGAG TGAATGTGT ATATTATTTG TGATCATAGC ATTAGTTCTT
60801 TTTCTGCTAT ACCCTGCATA TCTTCAAAGT CACAGTGTGT GTCCTGCCAT
60851 CTCATTAGTG AATTGTACCT AGATTATGTG TGTGCCCCCT TTGTATGATG
60901 TTTCTGGAAC GCTATAAGCA GCTTTTAGAG TCAAATGCAT TCATTTTAAAC
60951 TGGCTTTATG TCTAGTGGT TTCATGACTA CAAATTTGAA TTATCTTACT
61001 GCATAACATA AAAAATGTCT GGCTTTAGCA ATTAATGCCC GAAATTATTT
61051 TGCCCTGCAA TTGTCATACC TGTATGAAAC CTGTCCCAGT TTGCTTAAGT
61101 GCACAACTGA TTATGTATTC CTGTGTGTAT GCTAATATTT CACAAGTGT
61151 TCATGCATCC TTTTTTAAAA AACTACTAAC CAGAATATTA TCGTAGCTAC
61201 TCATTCATTC TGCTTTCTGC TTCACCTATA ATAATCTTTT AGGACTGCCT
61251 TCTGATTTTT CACCTATCTT TTAATGTAAG CATTAACAAC TAAGACTTTC
61301 ATAAAAGCAC TATATCTTAA CTTTCCTGGC CTAAATCAAA AAAAGGAAAA
61351 CATTGATAAG TGTCTAGAA ACTTGGATTC TTTTATAGAT TTGTTCTTGG
61401 GGCTCTGATG TTTGGGATTG ACGTCTCTGT CTGACCATTT TATATGCATT
61451 TTATCTTAAT AGTATGTGCT TTCATGAAGA TTCTGATACA AGTGGGCAAT
61501 CCTTAAATTA TCTTTGAAAA ATTGGTTAAT TTTGGTTAAA AAAGGGAAAG
61551 TGCTGGGTG CAGTGGCTCA CGCCTGTAAT CCCAGCACT TTGGGAGGCC
61601 GGGACGGGTG GATCACAAGG TCAGGAGTTG AAGCCCATTC TGGCCAACAT
61651 GGTGAAACCC TGTCTCTACT GAAAATAATT GGGGCATGGT GGCACATGCC
61701 TGTAATCCCA GCTACTTGGG AAGCTGAGGC AGGAGAATTG CTTGAACCGG
61751 GGACCAGGA GCGGAGGTT GCAGTGAGCT GAGATCGCGC CACTGCACTC
61801 CAGCTGGGC TACAGAGCGA GACTCTGTCT CAAAAATAA ATAAATAAAT
61851 AAATGAAAAA GAGAAAATAT TGAGAGGATT TGGTCATCAT TTTACTGCTC
61901 TCTTCATGTG ATGGAAATCA ATTTCCCTTC TCAAATGGGA TCAGTATCAT
61951 TTCTTAGTCA TACATCCATC CAGTTTTTGT TACTTTTTTG TTGGCATACA
62001 TTAATCAAAA TAGCTCTGCT TCATTGAGGC ATGCAGTCCT CAGACTCTCG
62051 GTGGAAAGGC TGTCATACTA TTAGTGACCA TAGTAACTTT TTATACCAAA
62101 GGATGGTTGC TGATAATTT TAATATCTTT ACCAATAAAG TACTTTTTTG
62151 AAATACAAAA TCAGGCTGCT TGCTTTGCTC TATTCCTGTC AACAAAAAGG
62201 ATTTAGCTAT AGATTTAGCT TCTCCTTTTA TTTTCCCTTT TATTTCATAG
62251 GAGTCTTCTG TTTATTCCTT TCAGGCGCCT CCTTGGCATT ATAACAAAAA
62301 AAGATATCCT CCGCATATG GCCAGACGG CAAACCAAGA CCCCCTTCA
62351 ATAATGTTCA ACTGAATCTC ACAGATGAGG AGAGAGAAGA AACGGAAGAG
62401 GAAGTTTATT TGTGAATAG CACAACCTT TAACCTGAGG GAGTCATCTA
62451 CTTTTTTTTT CTCCTTTACA AAAAAAGAAA GGAAATATAA AAGCCGGGTT
62501 TTTGCAACAT GGTTTGCAAA TAATGCTGGT GGAATGGAGG AGTTGTTTGG
62551 GGAGGGAAG GAGAGAGAAG GAAAGGAGTG AGGTATTTC CGTCTAACAG
62601 AAAGCAGCGT ATCAACTCCT ATTGTTCTGC ACTGGATGCA TTCAGCTGAG
62651 GATGTGCCCT ATAGTCAGG CTGCGCCTC AACAGAGATG ACAGCAGAGT
62701 CCTCGAGCAC CTGGCCTGTT GCTCCAACAT TGCAAAGACA CATTATCAGT
62751 CCCTATTTCT AGAGGGATTA CTTTGAATTG AGCCATCTAT AAAACTGCAA
62801 GGTCTTGCCC TTTTTTTTAA TCAAACTGT TCTGTTTAAAT TCATGAATTG
62851 TATAGTTAAG CATTACCTTT CTACATTCCA GAAGAGCCTT TATTTCTCTC
62901 TCTCTCTCTC TCTCTCTCTC TCTCTCTACT GAGCTGTAAAC AAAGCCTCTT
62951 TAAATCGGTG TATCCTTTTG AAGCAGTCCT TTCTCATATT GAGATGTACT

FIGURE 3, page 20 of 27

63001 GTGATTTTAC TGAGGTTTCA TCACAAGAAG GGAGTGTTTC TTGTGCCATT
63051 AACCATGTAG TTTGTACCAT CACTAAATGC TTGGAACAGT ACACATGCAC
63101 CACAACAAAG GCTCATCAAA CAGGTAAAGT CTCGAAGGAA GCGAGAACGA
63151 AATCTCTCAT TGTGTGCCGT GTGGCTCAAA ACCGAAAACA ATGAAGCTTG
63201 GTTTTAAAGG ATAAAGTTTT CTTTTTTGTT TTCCTCTCAG ACTTTATGGA
63251 TAATGTGACC GGGTCTTATG CAAATTTTCT ATTTCTAAAA CTACTACTAT
63301 GATATACAAG TGCTGTTGAG CATAATTAAA TAAAATGCTG CTGCTTTGAC
63351 AGTAAAGAGA AGGAAGTATT CTGATTAGCT GTATCTGGTA TTAATTGCAT
63401 GTTAAACAC TGGAATTTTT AAAATTGAAA TTAGATCAGT CATTCTTTTC
63451 TTTTCTCAAG ATATCTCATG GCTGACACTG AAGAAGAAAT GTAATTCATA
63501 ACTTGCACTA AATGTATATT TTTTTTCTTA AAAATTTACC ATTCTTATTT
63551 ATATTTTAT GGATTAAAT TTATAAAATA CAGATCAGTT AATATTGCAC
63601 TTAAGTAATT TTACCTTTTT AATGTGATTT TTATAGAATA ATTCAGACTT
63651 ACAAATACAG AGATATGAAC AAAGTTTACA GTGGGAACAA AGGTTTAAAA
63701 AAAGGTTGTG GTTCTCTCTC TGTGATCCAG TGTGCACATA AACCTTTCTC
63751 TGATCTTTCA CTGCCATCCT CTGGATTATG TCTTCTGACC TGTCCATTTT
63801 GACCCATTAA CTGGAAAGTT GAAAAACTAC ATTAAGTGAAGT AAGTTGAAAA
63851 ACTACATTAC TTTGGAGAAT AAAACCGAAA GTTCGTGTAT ACCTTCTTAA
63901 AAAAAAATC AAACCAAAAA TGTGAAAACA ATAGAATTGC AAAGATAGCA
63951 GTTAAAAATTT TAATCTGAAA ATAACCTTTG AATCTCGGGC TAGGTTATGT
64001 CCATATTTGA AGTGGTCAGT GATGGTTTGA ACATTTTTTG CAGGATGAGT
64051 TAAAATGCAC TGGATTATAT TTGGGATTTT TGTTTTTGA ATTGTCTGTT
64101 TTAATCACAG CCTTAATTCA CAATTGGCAA AGGCAGTTTA CTCAAAGGAC
64151 TGGGCTAAAT ATCTCTAAT TATGCATTTT TGATAGGAAA ATGAAATTTT
64201 TGCAAACAGA CATTTTCTTT TTTTTTGGCT GGAGTGCAGT GGGGCATGGT
64251 CTTGGCTCAC TGCAGCGTTG ACCACCTGGG CTCAAGTGAT ACTCCCGCCT
64301 CAGCCACCCA AGTAGCTGGC ACTACGGGCA CACGCCACCA TGCCCAGCTA
64351 ATTTTTTTGT ATTTTATAGTA GAGATGGGGT TTTGCCATGC TGCCCAGGCT
64401 GGTCTCAACT CCTCAGCTCA AGCAATCTGC CTGCGTGAGC CTCCCAAAGT
64451 GGTGGAATTA CAGGCGTGGG CCACTGCGCC TGGCCCAGAC AGACATTTTC
64501 TGAAACACAA CTGGCAATGA GCTGTTTTTA CATTTTGAAA GTGATTCTTC
64551 ACTTCCTAGT TCTTAATTAT AGTATACCTA TTAAGATCTG TAAGATCCTG
64601 AAGACATAAG ATCATGAAGC CATATAAGAA TGAGGATTGA AAGTTGAGCA
64651 AAATTTTTCGG GATTTTGGGA AACATTCTTA GCTGTGCTAT CTGCCTAAAA
64701 TTATTCCTTA TTACTTCTCT CCTTTGACAG ACTTCAAGTT TTCTTCATAG
64751 CCCTTTCAAA GTTTTTTGAG CCATCCAGAG TAAAATCATT TCTAAATGAT
64801 AGTTCTGTAT ATCTCCAAT CGTCTTAAGT GTATTTGCCT GTGTGCAACG
64851 TATTGCTAGA CTATGAATC CTCAGCATGG CTGCTGGATA ACTTAATTGT
64901 CCTGAGTTAA TAGCCTTCAA AGGACAAATC GGTTCCTTTG CAGATAGCTT
64951 CGTAAAACCT CACATGGAGT TTATTTTATC ATATTTCCCT TTTTATTTTC
65001 TGCTCCTCCT TTAATTGCCC ATCTTGCTTC AGAGACTGAC ATTTCAGGGT
65051 GGATATTAAT TAAAGCATT ATTTTGTTTT TTGGTATATT TCTATCCCTA
65101 GTATTTCTAT CTTACTGCTA AAATACAGGA AAAGTGCCGT ATTTTAAATG
65151 CATTTAGTGG TTTTCTTTGG TGTTATCTGT TCCATTTTTC TTTTTCATAC
65201 ATTGAAGTGT GTCTCCTTTT CAACCAAAAT AATGAAATAG TGGAGACCAT
65251 GAAATTGTTG TGCCTGGCTA ATTGGCAAAT TAATTTACCA ATATAATAAG
65301 TGTAGCGCCT TGTTTGAATA CCCTTTTGA GAAGGTATGA TGAGAATGGG
65351 CAAGGGTGT (SEQ ID NO:3)

FEATURES :

Start: 2159
Exon: 2159-2237
Intron: 2238-22041
Exon: 22042-22199
Intron: 22200-30359
Exon: 30360-30459
Intron: 30460-31475
Exon: 31476-31663
Intron: 31664-32964
Exon: 32965-33087
Intron: 33088-34548
Exon: 34549-34755

Intron: 34756-37975
 Exon: 37976-38056
 Intron: 38057-39552
 Exon: 39553-40098
 Intron: 40099-46366
 Exon: 46367-46553
 Intron: 46554-49237
 Exon: 49238-49636
 Intron: 49637-55445
 Exon: 55446-55662
 Intron: 55663-62274
 Exon: 62275-62362
 Stop: 62363

CHROMOSOME MAP POSITION:

Chromosome 4

ALLELIC VARIANTS (SNPs):

DNA			
Position	Major	Minor	Domain
1275	T	C	Beyond ORF(5')
1456	T	C	Beyond ORF(5')
5893	G	A	Intron
6226	A	G	Intron
8866	T	G	Intron
10397	C	T	Intron
10621	T	-	Intron
19651	A	G T	Intron
19891	T	-	Intron
20272	C	A	Intron
20412	T	A	Intron
23340	A	G	Intron
29948	T	A	Intron
33579	A	C	Intron
40762	G	A	Intron
40936	T	C	Intron
45998	A	G	Intron
47771	T	C	Intron
48117	C	T	Intron
54563	T	G	Intron
58735	C	T	Intron
59643	C	A	Intron
61638	G	T	Intron
63291	G	C	Beyond ORF(3')
63463	A	G	Beyond ORF(3')
63636	G	A	Beyond ORF(3')
63998	T	C	Beyond ORF(3')

Context:

DNA

Position
 1275

GCATTTTCAGGAGGAGAATCTCCAGTCTAGAGGAATCCTCTCAGAGGTAGCTATAAAATA
 TTGAACTCTGATCTTCAATAAGCATTGTGCGGTTTTTGTGTTTTTAAATGACAGTTT
 TAAACAAGAAAGTTGCTTTATTTCTGAACCTCATAAAAATTTCTATTAAAGAGACAATTT
 CTGAATTTTATAACAATTTCTAGAACAGTTGAGTACCTCACTTTGAGACACATTTTGCT
 AAAAGTTAAAAACACAAAACCTTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAG
 [T,C]
 CCTCTGCTTAGTAAACCTCTTTTTTGCCTAGTTTAGACACATAACAATAGTAAAGTTACTT
 AGTACGTTGATAGTTTTCTTTCTCCTCAAAAGCTACAATGTCTTACTAGCTAGTTCCTTC
 AAGAAAGGAAACAAGAAGCCGCTGGAGGAGATTGGTGAGTGGGATAAAACACTATTCAAC

TCTTCAGTTATTTCGGTTTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTTT
TTTTTATTTTTAATAGACTTAGAACCAAGTTTCTTGAGAAACCTTTGGCATATTGTAGTT

1456 TGAATTTTATAACAATTTCTAGAACAGTTGAGTACCTCACTTTGAGACACATTTTTGCTA
AAAGTTAAAAACACAAAACCTTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAGT
CCTCTGCTTAGTAAACCTCTTTTTTGCCTAGTTTAGACACATACATAGTAAAGTTACTT
AGTACGTTGATAGTTTTCTTCTCCTCAAAAGCTACAATGTCTTACTAGCTAGTTCCTTC
AAGAAAGGAAACAAGAAGCCGCTGGAGGAGATTGGTGAGTGGGATAAAACACTATTCAAC
[T, C]
CTTCAGTTATTTCGGTTTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTTT
TTTTTATTTTTAATAGACTTAGAACCAAGTTTCTTGAGAAACCTTTGGCATATTGTAGTTT
TTTTATGGCTATGACTCACATGACATTACTGTATAAACTAGTACATTCTCTCGTAAAC
CACACAACTTACTAGAGTGTGCTCTCATTTTTCTACATTAGAAATGAAAAGGGCATT
GTCTGCATTCAAATTTCTTTTTACATCTCTGTATTACTTTTTCCCTTTATATTTATC

5893 TCTAGTTGACAAGACTGAGGTAAGGAATTGTTAAGGAAAAGTCAGAATTCATCCAGATA
TTTGGCTCATACTTTAATCATGAGGCTAAACTGCTTCTCTACACGTATCTTCATAGTA
ACTTGTGTTTTAAGTCTGGTAGAAGCATAAGAAGTTTAAACACAGACAGAATCCTGTGGA
AGTTAGTAAATTTCTAGTGAACGATAGAAATGATAGAAATCTCTTCTCCCCCAAAGTCC
CAAGAACAGATTAGTCTGCTTTTGACAAGTGTTATCAAAGTAGACTGTCTCACATACAC
[G, A]
GGGGACTCAATAGGGCATTCTGGTGGATATAATAAAATGAGTAAATGCGATAACAGGAG
GAAATGCCTAGTGTGTTGCTCTTGGATTAGTTTTGATACAACAAAGGCAGCTTTGTTGTG
AGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGAAGTTGGAAGAGTGGCAGATCCTAGAGGA
CTAATGATGGGCTTAAACCACAAAAGTGTGCTTTGCCATTGAA

6226 ATAAATGAGTAAATGCGATAACAGGAGGAAATGCCTAGTGTGTTGCTCTTGGATTAGTT
TTGATACAACAAAGGCAGCTTTGTTGTGAGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGA
AGTTGGAAGAGTGGCAGATCCTAGAGGACTAATGATGGGCTTAAACCACAAAAGTGTG
CTTTGCCATTGAAATAAAAGTTTGGGGTCTTATTTTTCAATTTCTCCCTGAAATTATT
TCTTGACATTCATTAGCTCAGCAGTGTATCTAAATAAAGCTTTTTTGGGTTTCTATTATA
[A, G]
TAGAGGTTTGTTCCTTTTTCTTCCCTTTGAAAAGTATCATTTTTTGCACATTATTTGAAA
ATCCAGTGTTATATGATATTCTTATTGCCAGAGGGACATTCTGCAGGCTCTTTGTA
TGATTTTAGGATTAGATACTTATTATATTTTTATTGGCCCTAATATTTTATCCAAC
AAAATTAACCTCTTCTTAAATAATCCATCTAAGTGTCTGTAAATTAAGGAACAC
TAAAGATTCTTTATTTGGTGTGAGAACTCCTTGTCTACAACAGTAGTATAAAACAAA

8866 ACATGTAAACCAACAATGAAATTATTTTAGTGACTTGAGAATCAAAGTGCTAGAGTTTGA
ATCCCTGTTCTACTACTTGCTAGCGGTGTGACCTTGGGCCTGTTTAACTCTTGACACCT
GTTTTCCAAATTTATAAAGTGGAGATAATAATATCTGTACATTGTGTTGTTGTGAGGAT
TATATGAACATAATATGTAATGTCCTGAGAACAATGTCTGGTACACATTAGTTAATTA
AAATTAGCTGTTCTTACTGTTATTATTAGACATGAGCTAGATAACAGTGGCCTCTACATG
[T, G]
GAAAGATTATTTTAAATCTGATGTAGTTTCACTTATCTATTTTTTTTATTTTTGTCCCTT
TTGCATTGATGTCATATCTAAAAAACCTGCCTAACTCAGGATCACAAAATTTACTCCTG
TATTTTATAATTTTAGCTCTTTAGATCTAGGATCCATTTTATAGCTAATTTTATATATGG
TGTGAGGTAGGGGTACGGTTTCTTCTTGCACGTGAATAGCCAGTTGTCCAGCATCA
TTTATTCAAAGACTATTCTTCTCTACTAGAAAAATATTTCTTTAAAGAATAATGAAT

10397 CCAGGCTCCCTTGAACCTCTGGGCTCAGATGATATAGCCTCCTGCCACAGCGTCTGATT
AGCTGGGACTACAGGTGTGCACCCTACACGTGGCTTCTCTGATGAAATTTTAAATACCC
AAATATTTGAGCAGAAATATAGCTTGTGTTTATTTGTTTTCTACTATCTGTCAAGTATA
GTATTAAATGTTTTACATAATTTGTCTCCAGTCCACATACAATACTCTAGTAGAAGTGGG
TAACAAAACCAAGGTACTCAAAGAGGTTAATAAGTAACTTGGCTGGATCACAGAATAA
[C, T]
GGGAGGCAGGGCTGGAATTTGACTCTAGGTCTTTCTGACCTCAAAGTGCAGTAAAGTCAT
GGAATTTCTCTACTAGGCCACCTGGAAGAAAAGTGTCTTTTTTCCAGTCTTTTTGTTA
CTGTTTTTTCAGCCAGGAGATAGTAGAGTTAGGTAGTAGAATAGTAGTCACTGGCATCCGG
TAGTCAGCCCTCCAAAAAGTTTTTGTATTTTTTTTTTTTTTTTGTCTTAACTTGAAG
CTACTAACTTTTCAGGTACACTTTCTTATCATCCAAGAGCTGGATATTTAGGTAGCAGAA

10621 CTCTAGTAGAAGTGGGTAACAAAACCAAGGTACTCAAAGAGGTTAATAAGTAACTTGCGC

FIGURE 3, page 23 of 27

TGGATCACAGAACTAACGGGAGGCAGGGCTGGAATTTGACTCTAGGTCTTTCTGACCTCA
AAGTGCAGTAAAGTCATGGAATTTCTCTACTAGGCCACCTGGAAGAAAAGTGATCTTTTT
TCCAGTCTTTTTTGTACTGTTTTTCAGCCAGGAGATAGTAGAGTTAGGTAGTAGAATAG
TAGTCACTGGCATCCGGTAGTCAGCCCTCCAAAAAGTTTTTGATTTTTTTTTTTTTTTT
[T, -]
GTCTTAAACTTGAAGCTACTAACTTTTCAGGTCATACCTTTCTTATCATCCAAGAGCTGGA
TATTTAGGTAGCAGAACTATGGAATTATCCTAAGTCCTCTTGAAGCTTCAGCTGTTAAA
ATTAATTGGTTCGATTAACTGTGCTCAAGATTTACATTTCTAGGAGCCACAGTTTGA
TTGGTCTAACTTGATCTATGTGTTTTCTTTAGCTGGGAGGAGAAGGTATCTTGATTGA
TACCTTCACCAGGACTGCATGCAGTGAGGGACAGAAGTTTCCTTAAAAATAATTGGGTCT

19651 TTTATTTTCTGCTACTATGGCAGAATTGAGTTGTTGCACTGTGTGGCATCCAAAGCCTA
AAATATTTACTCTCCTGGCTCTTTGCCAACCGTTTTAGATTATGAGCACTTTGGCATT
TTATGTTTTTGTCTTTCTTATAGCACACAGTAAGATGTTCTGCCACATTGTGCATAA
TTTATGGGTTTATTCAAGGATTTATGCAAGGTAGCTGCAAGAAAAAACCTAGAAGTGA
ACTTGCTAGGTTGAAGAGCA
[A, G, T]
CTGTGTATGTTAAATTTTGTAGCTTTTCGCTTCCCAAAGGGATTATTCATTTCTACT
TAACTACTAATTTGTGATAGGACTTCTTTCTCCATAGCTTTGCTAAATTAATGCATTC
ACACACTTCATCTTTACTAATCTGATAGAGGAAATGATATTGTGGATTGATTTGCATT
TCTTTTTATGTGTAGCTTGAGCTTATTTTCATATTTAAAGCCAATTGTATTTCTTTT
CTTGAGCTATCTTTAATGT

19891 TTTATGCAAGTGTAGCTGCAAGAAAAAACCTAGAAGTGAAGTGTAGGTTGAAGAGCA
TCTGTGTATGTTAAATTTTGTAGCTTTTCGCTTCCCAAAGGGATTATTCATTTCTACT
TTAACTACTAATTTGTGATAGGACTTCTTTCTCCATAGCTTTGCTAAATTAATGCATT
CACACACTTCATCTTTACTAATCTGATAGAGGAAATGATATTGTGGATTGATTTGCAT
TTCTTTTTATGTGTAGCTTGAGCTTATTTTCATATTTAAAGCCAATTGTATTTCTTTT
[T, -]
CTTGAGCTATCTTTAATGTCTTCTGATACATTTCTGAAGTCTGTGATACTCATATAA
GATATATGGTGAACATGTGTCAAAGATTTATTTGACTCTAATGAGGGAACCCGCTGATG
ACAAGGCTGATTGAGAAGAGGATGTGTGAGATGAAGTGTATATCATCAGTGAAAGAAAGC
AAATTTCTACAGGGCAAAAACAAAACCAACTCTAAGGGTTATTGTTTCTACTGGACAG
AATTCATTTGCATTTTACCAGATAAAAAATTACTATTTTCAATTTATCTTTTACAAATCAT

20272 CAAAGATTTATTTGACTCTAATGAGGGAACCCGCTGATGACAAGGCTGATTGAGAAGAG
GATGTGTGAGATGAAGTGTATATCATCAGTGAAAGAAAGCAAATTTCTACAGGGCAAAAA
CAAAACCAACTCTAAGGGTTATTGTTTCTACTGGACAGAATTCATTTGCATTTTACCA
GATAAAATTAATTTTCAATTTATCTTTTACAAATCATTTTCTAATTTTACAGAGTCT
ATTCCCTAATCAGTAGTAAATAGTCTTCAAATTTCTCCGACGCTCAGGTGACTATTATG
[C, A]
AGGCTAATTTGTTGACACTCGGGCTTGACTTTAAGAGAACATGCCATAATCTTTTGGCCTT
ACTTCCAAGTTTTGGATAATTTTTCTTAACACATTTTTCTCTAATTGCAATGATTTCAAG
TGATATTATTTCTTTTTTTTAAATTTTTTACTATTTATTGATCACTCTTGGGTGTTTCT
CGGAGAGGGGATTGTCAGGGTCATAGGACAATAGTGGAGGGAAGGTGAGCAGATAAAC
ATGTGAACAAAGGTCTCTGGTTTTCTAGGCAGAGGACCTGCGGCCTTCCACAGTGTTT

20412 TTATTGTTTCTACTGGACAGAATTCATTTGCATTTTACCAGATAAAAAATTACTATTTTCA
ATTTATCTTTTACAAATCATTTTCTAATTTTACAGAGTCTATTCCCTAATCAGTAGTAAA
TAGTCTTCAAAATTTCTCCGACGCTCAGGTGACTATTATGCAGGCTAATTGTTGACACTC
GGGCTTGACTTTAAGAGAACATGCCATAATCTTTTGGCCTTACTTCCAAGTTTGGATAA
TTTTTCTTAACACATTTTCTCTAATTGCAATGATTTCAAGTGATATTATTTCTTTTTT
[T, A]
AAATTTTTTTTACTATTTATTGATCACTCTTGGGTGTTTCTCGGAGAGGGGGATTGCGAG
GGTCATAGGACAATAGTGGAGGGAAGGTGAGCAGATAAACATGTGAACAAAGGTCTCTGG
TTTTCTAGGCAGAGGACCTTCCGACGCTTCCACAGTGTTTGTGTCCTGGGTACTTGAGA
TTAGGGAGTGGTGATGACTCTTAATGAGCATGCTGCCTTCAAGCATCTGTTTAAACAAAGC
ACATCTTGACCGCCCTTAATCCCTTAAACCCTGAGTTGACATAGCACATGTTTCAGAGA

23340 TTTTTTTTTTGGAGGTCGGGGGACTGTGCCCCATTCTGTTGCCCAAACCTGGAGTGCAAGT
GTGCAATCTTGGCTCACTGCAACCTCTGCCTCCCAGGTTCAAGCGATTCTTGTACTCAGC
CTCCTGAGTAGCTGGAATTATAGGTGTGTGCCATCATGCCAAGCTAATTTTTGTATTTTT
AGTAGAGATGAAGTTTCGCCATGTTGGCGAGGCTAGTCTCAGACTCCTGGCCTCAAGTGA

FIGURE 3, page 24 of 27

29948

33579

40762

40936

45998

FIGURE 3, page 25 of 27

TTTTGTAAGTGAATTCATGGGATTGTTAAGGTACAAGATTTTGCTTTAGTTTTATTTGTA
CTAGGATTTTGCTATATTAATAACAATGTGAAAAGAATCAAAAGTGTTAGAAATAAATGCA

47771 GAAGAGTAGAACATGAGGCTTTATTTTAAAAGATTAGCAGAATTTAAGGAAAAGGTGACTT
TGTTGAAGATTATAATGTGAAGACAAAGGAACGAGGATGGGAATAAATTTTGTATTCATG
AGGCTTTGAAGAAATTGACTCTAGAGAGTATATTTTGGGTACTTTTGGGAAATGAAGTTG
GATTAGTGAGAAGGAACAGATTATGAAAAGACAAGAAACCTGATTAATGTCAGGATGATT
TTATATTTGAAG
[T, C]
TGGTCAGATTTTATGGCAGTCCTGGCTTTGCCATTTTGTAGTTTGATGACTTTGAGAAAAGTT
CCTTCTTGAAGTTTAAATTTTCTGTATATAAAAAGTAATAACACCTGGTGATCTGCTAGG
TTTGTTTTGAAGATTATATGAGATAAAATGCATGCAAACTGTTATAATAGTGCCTGGTA
AAATAAGTGCCTAGTTTTTAAAAACAAGTCTTTGTAACTGCTTAGGACATGCCTGGTATA
GGGTAGGTATGT

48117 GACTTTGAGAAAAGTTCCTTCTTGAAGTTTAAATTTTCTGTATATAAAAAGTAATAACACC
TGGTGATCTGCTAGGTTTGTCTTGGAGATTATATGAGATAAAATGCATGCAAACTGTTA
TAATAGTGCCTGGTAAATAAGTGCCTAGTTTTTAAAAACAAGTCTTTGTAACTGCTTAG
GACATGCCTGGTATAGGGTAGGTATGTAATACATAGTAGGTAGGATCTGTCTCCTTGCTA
TTTTTAGGTAAAAAACAAGGAAGAGCTTCAGCTTAATACAGTATGAACTGACGAGCC
[C, T]
TGGTAGGTTTTTGAAGCAAAAGAGCAACACAGTAAAGTAGTACTTAGGAAAAGATTACAA
GGGAACATGGCTTATACAGTGGTAATGGGGCTGGAGTCAAGGAGGTAAGATAAAATGGT
ATTATAAATTAAGGAATAGCCAGGCACGATGGCACATGCATGTAATGCCAGCTACTGGAGA
GGCTGAGGTGGGAGGATCATGGGAGTCCAGGAGTTTGTAGACCAGCCTGGGCAACTGAGTG
AGACCCCAAATCCTAAAAAATACAAAGTAAAAAGGAATAAAGTCATGAGGGCTTGGACT

54563 GCTTTGTCAACCCAGGCTGGAGTGTGGTGGTATGATCATGGCTGACTGCAGCCCTGACCTT
CCGGGCTCAAGTGATCTTTCCACCTCAGCCTCCCAATTACTTGGGACCACCAGCATGCTT
GGCCGATTTTTTTTTTTTTTTTTTTTGTAGAAGCAAGGTTTTCCCTATGTTGCCAAGGC
TGGTCTTGAACCTTTAGGGCTCATGTGATACTCCTGCCTCGGCCCTCCCAAAGTGTTAGGAT
TACAAGCCTGAGCCACCATGGCCGCCAAAATATTTTCACTATAACAAATATCATATCTG
[T, G]
ATATACTCAGTTTTTAATACTAACTCAAAGTAGAAACATAAAGCTGAATGACTATTTTATT
TTCAGATTCTCTCCATTGAGTTTCCTTCTCCGTCTTGTGTGATCTCTGAACTTTTCTCCA
TCTTTGCCACTTCTGTCTAGCATTTTTTTTTTATCAGCAGTTTCATTGAGATTTTTTTT
TTAGTTCTTTCAACGGTGGAGTGGAGTGAAGTAGGCAGCAGGACAGAAGACTTGAAGCAGAGC
ACACTGGAGAGGAGAAATTAACAAAGCCTTTATGAATAAAACAACCCCCCAATATCAGTC

58735 TGGGTTATGCCCTGTAACTCTTACATCATTAGTTTTTAGCCCAAAGGAAACAGCAAAAT
AATGTTTTATATGAGCCACATTTTGCCTTGATTTTCTTCCACTCTGTAAATTAATAAA
GCAGCACTCTGACTTTATTATGCTCAAATCGCTCTTCCATTAAATGTGTGTTTCTCCAT
CTTTTAGGTTTTTTTACTTTTATAAATACAGAGATTACTGTGTAATAATCTAAATTTGCCAC
TGGGTCGTTATACATTTGTAACCTTCTCACAGTATATTTTGTGATTGGCAGAGTTTAC
[C, T]
AATATAGATGATACTAACTGAAATTAATCATTCTGTATAATTGGATAGAAAAGCATGAGT
AAGAATTCATTTGGTATTATATTTAATTAATTGCCAAGATTTTACATTTCTGACTACA
ACAATAAAATCAAATGAATTGATGGCTTAAAAAAGAAATCTCAAATGTTAGTCAATG
AAGAACATCTATTGAATGAGTGAATGTTTATTATATATAGTGCATTTTCTGAGCTTTTTT
GGAGGGGAAGTTGCTCCATGCTCTGAGAACTTTAAGGATCGATACATTATTTTTTAAC

59643 GTTTATATTGCCACATTAATTTCCATTATAAAACCAGTAACCATAGTTTTGTTTTAATTA
GCAATCTAATTATTTTTCATGTATCCTCATTATGAGAATTTATGTCCATCACTTTGCTTGA
TGTGATAACAGTGACATGCTAAATGAGAAACAATTGTTATTTAGAAAAAATGCACAAAG
TGAAAGTCTTTTAAATCCCTAATCATAAATACATTTTATTAGCTTACTTTAAGAAGTGGC
AGTCACAGCTCCTGAACATTAGGGAGTGTCTTTTGGTCAGCATTATTTATTTAGTGCA
[C, A]
ATTGCCTTTAATTTTAAATTTGAAATTATAGTAAATCCACGGGAGTTTTTAAAGTCTCCTC
ACAGCCTTTTGTACCTTTTCCCAAGGTAGATCCAGATGATAACTGCTGTGTGTGACA
TCATAGAAATTAGAAAAATATTTTCTCTGAGGAAAGAACATTGTAAATGAACTCTACA
TATCAGAGGTCTATAGCTATGTATCAATATTAAGTTTCTTTGTACTTTGCTTTGTAGTC
ATCTTCATTCCAACTTTTATAATTATTTTTTACTTTTAAAAAGAAAAATAACCCACCA

61638 AAAAAAAGGAAAACATTGATAAGTGTCTAGAACTTGGATTCTTTTATAGATTTGTTCT
TGGGGCTCTGATGTTTGGGATTGACGTTCTGTGCTGACCATTTTATATGCATTTTATCTT
AATAGTATGTGCTTTTCATGAAGATTCTGATACAAGTGGGCAATCCTTAAATTATCTTTGA
AAAATTGGTTAATTTTGGTTAAAAAAGGAAAGTGGCTGGGTGCAGTGGCTCACGCCTGT
AATCCCCAGCACTTTGGGAGGCCGGGACGGGTGGATCACAAGGTGAGGAGTTGAAGCCCA
[G, T]
TCTGGCCAACATGGTGAAACCCTGTCTCTACTGAAAATAATTGGGGCATGGTGGCACATG
CCTGTAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATTGCTTGAACCGGGGACCCAG
GAGGCGGAGGTTGCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGCTACAGAGC
GAGACTCTGTCTCAAAAAATAAATAAATAAATAAATGAAAAAGAGAAAATATTGAGAGGA
TTTGGTCATCATTTTACTGCTCTCTTCATGTGATGAAATCAATTTTCCTTCTCAATGG

63291 GAGATGTACTGTGATTTTACTGAGGTTTCATCACAAGAAGGGAGTGTTCCTTGTCGCATT
AACCATGTAGTTGTACCATCACTAAATGCTTGAACAGTACACATGCACCACAACAAAG
GCTCATCAACAGGTAAGTCTCGAAGGAAGCGAGAACGAAATCTCTCATTGTGTGCCGT
GTGGCTCAAAACCGAAAACAATGAAGCTTGGTTTTTAAAGGATAAAGTTTTCTTTTTGT
TTCTCTCAGACTTTATGGATAATGTGACCGGTCTTATGCAAATTTCTATTCTAAAA
[G, C]
TACTACTATGATATACAAGTGTGTTGAGCATAATTAATAAAAATGCTGCTGCTTTGACA
GTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAATTGCATGTTAAACACT
GGAATTTTTAAATTTGAAATTAGATCAGTCATTCTTTCTTTCTCAAGATATCTCATGG
CTGACACTGAAGAAGAAATGTAATTCATAACTGCACTAAATGTATATTTTTTTCTTAA
AAATTTACCATTCTTATTTATATTTTATGGATTAAATTTATAAAATACAGATCAGTTA

63463 TGTGCCGTGTGGCTCAAAACCGAAAACAATGAAGCTTGGTTTTAAAGGATAAAGTTTTCT
TTTTTGTTCCTCTCAGACTTTATGGATAATGTGACCGGTCTTATGCAAATTTCTAT
TTCTAAACTACTACTATGATATACAAGTGTGTTGAGCATAATTAATAAAAATGCTGCT
GCTTTGACAGTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAATTGCATGT
TAAACACTGGAATTTTAAATTTGAAATTAGATCAGTCATTCTTTCTTTCTCAAGAT
[A, G]
TCTCATGGCTGACACTGAAGAAGAAATGTAATTCATAACTGCACTAAATGTATATTTT
TTTCTTAAAAATTTACCATTCTTATTTATATTTTATGGATTAAAATTTATAAAATACAG
ATCAGTTAATATTGCACTTAAGTAATTTTACCTTTTAAATGTGATTTTTATAGAATAATT
CAGACTTACAAATACAGAGATATGAACAAAGTTTACAGTGGGAACAAAGGTTTAAAAAA
GGTTGTGGTTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCTTTCAGTG

63636 TGCTGCTGCTTTGACAGTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAAT
TGCATGTTAAACACTGGAATTTTTTAAATTTGAAATTAGATCAGTCATTCTTTCTTTTC
TCAAGATATCTCATGGCTGACACTGAAGAAGAAATGTAATTCATAACTGCACTAAATGT
ATATTTTTTTTCTTAAAAATTTACCATTCTTATTTATATTTTTATGGATTAAAAATTTATA
AAATACAGATCAGTTAATATTGCACTTAAGTAATTTTACCTTTTAAATGTGATTTTTATA
[G, A]
AATAATTACAGCTTACAAATACAGAGATATGAACAAAGTTTACAGTGGGAACAAAGTTT
AAAAAAGGTTGTGGTTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCT
TTCAGTCCATCCTCTGGATTATGTCTTCTGACCTGTCCATTTTGACCCATTAAGTGGAA
AGTTGAAAACTACATTAAGTGGAAAGTTGAAAACTACATTAAGTTTGGAGAATAAAACC
GAAAGTTCGTGTATACCTTCTTAAAAAAAATCAAACCAAAATGTGAAAACAATAGAA

63998 AAAAAAGGTTGTGGTTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCTT
TCACTGCCATCCTCTGGATTATGTCTTCTGACCTGTCCATTTTGACCCATTAAGTGGAAA
GTTGAAAACTACATTAAGTGGAAAGTTGAAAACTACATTACTTTGGAGAATAAAACCG
AAAGTTCGTGTATACCTTCTTAAAAAAAATCAAACCAAAATGTGAAAACAATAGAAAT
TGCAAGATAGCAGTTAAATTTTAACTGAAAATAACCTTTGAATCTCGGGCTAGGTTA
[T, C]
GTCCATATTTGAAGTGGTCACTGATGGTTTGAACATTTTTTGCAGGATGAGTTAAAATGC
ACTGGATTATTTGGGATTTTTGTTTTTGAATTTGCTGTTTTAATCACAGCCTTAATT
CACAATTGGCAAAGGCAGTTTACTCAAAGGACTGGGCTAAATATTTCTGTAATTATGCATT
TTTGATAGGAAAATGAAATTTTGAACAGACATTTCTTTTTTTTGGCTGGAGTGCA
GTGGGGCATGGTCTTGGCTCACTGCAGCGTTGACCCTGGGCTCAAGTGATACTCCCGC